

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 07:55:45 ; Search time 188 Seconds
(without alignments)
535.201 Million cell updates/sec

Title: US-09-942-052a-728

Perfect score: 1198

Sequence: 1 MAAPQLRRSRRCATPPRGDF.....LKSLMKLILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*
- 9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	4	AAM40223 Human pol
2	1198	100.0	229	5	ABG34856 Human can
3	1198	100.0	221	4	AAM42009 Human pol
4	138.5	11.6	164	5	ABG34855 Human pol
5	113	9.4	233	4	AAM93668 Human pol
6	113	9.4	233	4	AAG93290 Human pol
7	113	9.4	233	7	ADJ69163 Human pol
8	113	9.4	233	8	ADL31517 Human pol
9	92	7.7	695	5	ABBS4167 Human pol
10	92	7.7	695	8	ADS29336 Human pol
11	90	7.5	269	4	AAC08972 Human pol
12	89	7.4	572	8	ABM83354 Human pol
13	89	7.4	623	4	AAB93182 Human pol
14	89	7.4	623	5	ABBS97233 Human pol
15	89	7.4	652	4	AAB93168 Human pol
16	89	7.4	664	4	AAB83843 Human pol
17	89	7.4	664	4	AAB20219 Human pol
18	88.5	7.4	388	8	ADQ08702 Human pol
19	87.5	7.3	1245	7	ADC32818 Human pol
20	86.5	7.2	470	8	ABM83305 Human pol
21	86.5	7.2	669	6	ABU23878 Human pol
22	86	7.2	158	7	ADH88711 Human pol
23	85.5	7.1	571	8	ADU07910 Human pol
24	84.5	7.1	485	5	ABP51336 Human MDD

ALIGNMENTS

RESULT 1

AAM40223

XX AAM40223 standard; protein; 229 AA.

AC AAM40223;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3368.

DE Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI59379.

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Aab95788 Human pro
Adc31196 Human nov
Aau39678 Protonib
Abm36197 Protonib
Aaw59461 Microbial
Aaw59463 Microbial
Aaw59457 Microbial
Aaw59458 Microbial
Aaw59459 Microbial
Aaw59462 Microbial
Aaw59460 Microbial
Aaw59464 Microbial
Aaw59496 Microbial
Aap90110 Polypepti
Aaw59456 Microbial
Aaw59465 Microbial
Abu23749 Protein e
Adx72944 Plant ful
Aay90242 Human cyc
Adn72405 Thale cre
Aea16977 Arabidops

Example 5; SEQ ID NO 3368; 10078pp; English.

PS The invention relates to human nucleic acids (AA157799-AA161369) and the
XX encoded polypeptides (AA157799-AA161369) with nototropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification

SQ Sequence 229 AA;

Query Match 100.0%; Score 1198; DB 4; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.7e-127;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQPLRHSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGAGLGAEEP 60
DB 1 MAAQPLRHSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGAGLGAEEP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120

QY 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180
DB 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDOSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDOSKPEN 229

RESULT 2
ABG34856
ID ABG34856 standard; protein; 229 AA.
AC ABG34856;
XX
XX 15-JUL-2002 (first entry)
XX Human cancer related protein encoded by cDNA 85PIB3.
XX Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.
XX Homo sapiens.
XX WO200218578-A2.
XX
XX 07-MAR-2002.
XX
XX 28-AUG-2001; 2001WO-US026838.
XX
XX 28-AUG-2000; 2000US-0228432P.
XX (AGEN-) AGENSYS INC.
XX Raitano AB, Faris M, Hubert RS, Afar D, Ge W, Challita-Eid P;
XX Jakobovits A;
XX WPI; 2002-382963/41.
XX N-PSDB; ABK70506.
XX Composition for modulating the status of 85PIB3 protein or a molecule
XX comprising a substance e.g. antibody specific to, nucleic acid encoding,

PT or ribozyme of 85PIB3.

Claim 34; Fig 2; 201pp; English.

The invention relates to a composition comprising a substance that modulate the status of 85PIB3, where the status of a cell expresses 85PIB3 gene product is modulated. Also included are a composition comprising a peptide region of 5 amino acids of the 85PIB3 protein, in any whole number increment up to 229 that includes an aa position selected from an aa position having a value greater than 0.5 in the hydrophilicity profile, an aa position having a value less than 0.5 in the hydrophobicity profile, an aa position having a value greater than 0.5 in the percent accessible residue profile, an aa position having a value greater than 0.5 in the average flexibility profile, or an aa position having a value greater than 0.5 in the beta-turn profile; a polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous residues of the 85PIB3 protein; a recombinant protein comprising the antigen-binding region of a monoclonal antibody; a non-human transgenic animal that produces an antibody that binds to the 85PIB3 protein; a hybridoma that produces antibody specific to the protein; a single chain monoclonal antibody (MAB) that comprises the variable domains of the heavy and monoclonal antibodies specific to the protein; a vector comprising a polynucleotide that encodes the MAB; inhibiting growth of cancer cells or treating a patient who bears cancer cells that expresses the protein, by administering the protein, antibody, polynucleotide encoding the protein, antisense polynucleotide to the polynucleotide, ribozyme that cleaves the polynucleotide and T cells that specifically recognize the protein; and generating a mammalian immune response directed to the protein exposing cells of the mammal's immune system to an immunogenic portion of the protein or polynucleotide. The composition, which comprises an antibody specific to the protein, is useful for delivering a cytotoxic agent to a cell that expresses the protein by providing a cytotoxic agent conjugated to antibody and exposing the cell to the antibody-agent conjugate. The methods are useful for inhibiting growth of cancer cells or treating a patient who bears cancer cells that expresses the protein, for generating a mammalian immune response directed to the protein, for detecting the presence of the protein or polynucleotide in a biological sample in a patient who has or who is suspected of having cancer and for monitoring 85PIB3 in a biological sample from a patient who has or who is suspected of having cancer. The gene for 85PIB3 is located on human chromosome 15q14. The present sequence is the 85PIB3 protein

SQ Sequence 229 AA;

Query Match 100.0%; Score 1198; DB 5; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.7e-127;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQPLRHSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGAGLGAEEP 60
DB 1 MAAQPLRHSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGAGLGAEEP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120

QY 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180
DB 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDOSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDOSKPEN 229

RESULT 3
AAM42009
ID AAM42009 standard; protein; 231 AA.
XX
XX AAM42009;
XX
XX 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 6940.

XX DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-0052317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Zhou J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR N-PSDB; AA161165.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX Example 2; SEQ ID NO 6940; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the

CC encoded polypeptides (AA138642-AA142213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders. Note: The sequence data for this patent did not form

XX part of the printed specification

XX SQ Sequence 231 AA;

Query Match 100.0%; Score 1198; DB 4; Length 231;

Best Local Similarity 100.0%; Pred. No. 9.8e-127;

Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQAQLHRRRCATPPRGDFCGGTGTERAIDQASFTTSMEDWTQVVGKSSPLGAGLGAEEP 60

DB 3 MAQAQLHRRRCATPPRGDFCGGTGTERAIDQASFTTSMEDWTQVVGKSSPLGAGLGAEEP 62

QY 61 AAGPQLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVFSRVNTNVVLEAPFL 120

DB 63 AAGPQLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVFSRVNTNVVLEAPFL 122

QY 121 VGIEGSLKGYSTNLLFCGSCGIPVGFHYSTHAALARGHFCCLSSDKWVCYLLKTKAIV 180

DB 123 VGIEGSLKGYSTNLLFCGSCGIPVGFHYSTHAALARGHFCCLSSDKWVCYLLKTKAIV 182

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLSKLMKILSEVTDPQSKPEN 229

DB 183 NASEMDIQNVPLSEKIAELKEKIVLTHNRLSKLMKILSEVTDPQSKPEN 231

RESULT 4

ABG34855

ID ABG34855 standard; protein; 164 AA.

XX AC ABG34855;

XX DT 15-JUL-2002 (first entry)

XX DE Human cDNA 85PIB3 splice variant, open reading frame #3.

XX KW Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.

XX OS Homo sapiens.

XX PN WO200218578-A2.

XX PD 07-MAR-2002.

XX PF 28-AUG-2001; 2001WO-US026838.

XX PR 28-AUG-2000; 2000US-0228432P.

XX (AGEN-) AGENSYS INC.

XX Raitano AB, Paris M, Hubert RS, Afar D, Ge W, Challita-Eid P;

PI Jakobovits A;

DR WPI; 2002-382963/41.

DR N-PSDB; ABK70504.

XX Composition for modulating the status of 85PIB3 protein or a molecule

PT comprising a substance e.g. antibody specific to, nucleic acid encoding,

PT or ribozyme of 85PIB3.

XX Example 38; Page 124; 201pp; English.

XX The invention relates to a composition comprising a substance that

CC modulate the status of 85PIB3, where the status of a cell expresses

CC 85PIB3 gene product is modulated. Also included are a composition

CC comprising a peptide region of 5 amino acids of the 85PIB3 protein, in

CC any whole number increment up to 229 that includes an aa position

CC selected from an aa position having a value greater than 0.5 in the

CC hydrophilicity profile, an aa position having a value less than 0.5 in

CC the hydrophobicity profile, an aa position having a value greater than

CC 0.5 in the percent accessible residue profile, an aa position having a

CC value greater than 0.5 in the average flexibility profile, or an aa

CC position having a value greater than 0.5 in the beta-turn profile; a

CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous

CC residues of the 85PIB3 protein; a recombinant protein comprising the

CC antigen-binding region of a monoclonal antibody; a non-human transgenic

CC animal that produces an antibody that binds to the 85PIB3 protein; a

CC hybridoma that produces an antibody specific to the protein; a single chain

CC monoclonal antibody (Mab) that comprises the variable domains of the

CC heavy and monoclonal antibodies specific to the protein; a vector

CC comprising a polynucleotide that encodes the Mab; inhibiting growth of

CC cancer cells or treating a patient who bears cancer cells that expresses

CC the protein, by administering the protein, antibody, polynucleotide

CC encoding the protein, antisense polynucleotide to the polynucleotide,

CC ribozyme that cleaves the polynucleotide and T cells that specifically

CC recognize the protein; and generating a mammalian immune response

CC directed to the protein exposing cells of the mammal's immune system to

CC an immunogenic portion of the protein or polynucleotide. The composition,

CC which comprises an antibody specific to the protein, is useful for

CC delivering a cytotoxic agent to a cell that expresses the protein by

CC providing a cytotoxic agent conjugated to antibody and exposing the cell

CC to the antibody-agent conjugate. The methods are useful for inhibiting
CC growth of cancer cells or treating a patient who bears cancer cells that
CC expresses the protein, for generating a mammalian immune response
CC directed to the protein, for detecting the presence of the protein or
CC polynucleotide in a biological sample in a patient who has or who is
CC suspected of having cancer, and for monitoring 85PIB3 in a biological
CC sample from a patient who has or who is suspected of having cancer. The
CC gene for 85PIB3 is located on human chromosome 15q14. The present
CC sequence is the 85PIB3 protein
XX
SQ Sequence 164 AA;

Query Match 11.6%; Score 138.5; DB 5; Length 164;
Best Local Similarity 82.1%; Pred. No. 8.6e-07;
Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 162 FCLSSDRMVCYLLTKTKAIVNASEMDIONVPLSEKIAELK 200
Db 2 FFLSS-----YLLTKAIVNASEMDIONVPLSEKIAEVK 35

RESULT 5
AAM93668
ID AAM93668 standard; protein; 233 AA.

AC AAM93668;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3550.

KW Human; full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI; 2001-524255/58.

DR N-PSDB; AAK94603.

PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.

PS Claim 8; SEQ ID NO 3550; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO

XX Sequence 233 AA;

Query Match 9.4%; Score 113; DB 4; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0011;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
Qy 31 ASFTTSEWDTQVVKGSPLGPAIGAREPAAQPLPSWLQPERCAVFOCAQCHAVLADS 90
Db 49 ASMWSSSEDAV-----ADMERAGL--EEEAANAEE-----ERPLVFLCSGCRPLGDS 95
Qy 91 VHLAWDLR--SLGAVFVSRTNNVLRAPFLVGIKSLKSTYNNLLFCGSCGIPVGFHLY 149
Db 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKENGCVLETLCAGCSLNLGYVYR 153
Qy 150 STHAALAAALRGHPCILSSDKMVCYLL--KTKAIVNASEMDIONVPLSEKIAELKIVLTH 207
Db 154 CTPKNLDYKRDLCFLSVEAIESYVLGSSEKQIV--SEDKELFNL---BSRVEIKSLTQME 209
Qy 208 NRLKSLMKILSE 219
Db 210 DVLKALOMKLWE 221

RESULT 6

AAG93290

ID AAG93290 standard; protein; 233 AA.

AC AAG93290;

DT 13-SEP-2001 (first entry)

DE Human protein HP10650.

KW Human; gene therapy; tumour.

OS Homo sapiens.

PN WO200142302-A1.

PD 14-JUN-2001.

PF 06-DEC-2000; 2000WO-JP008631.

PR 06-DEC-1999; 99JP-00346863.

PR 06-DEC-1999; 99JP-00346864.

PR 08-FEB-2000; 2000JP-00031062.

PR 10-FEB-2000; 2000JP-00034090.

PR 10-FEB-2000; 2000JP-00034091.

PR 14-FEB-2000; 2000JP-00035829.

PR 14-FEB-2000; 2000JP-00035899.

PR 14-MAR-2000; 2000JP-00071161.

PR 30-MAY-2000; 2000JP-00160851.

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Kato S, Eguchi C, Saeki M;

DR WPI; 2001-381646/40.

DR N-PSDB; AAH68575.

PT Human protein originated from tumor cell line, applicable as drug,
PT reagent for studying intracellular protein networks and protein source
PT for drug screening, also encoded cDNA for gene diagnosis and gene
PT therapy.

PS Claim 1; Page 303-304; 471pp; Japanese.

XX The present sequence is a human protein. The human protein, preferably
CC originated from tumour cell line, is applicable as a drug, a reagent for
CC studying intracellular protein networks and a protein source for
CC screening proteins for binding low molecular weight drugs. The human
CC protein coding sequence is useful for gene diagnosis and gene therapy,
CC expression vectors and transformant cells for detection of ligands and
CC receptors

XX

present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, neurotropic, antidiabetic, anticonvulsant, antiarthritic, osteopathic, ophthalmological and cytoskeletal activities. This polypeptide sequence is a human heart mitochondrial protein of the invention.

Sequence 233 AA;

Query Match 9.4%; Score 113; DB 7; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0011;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTTSMEDTQVVGSSPLGAGLGAEPAPAGPOLPSWLQPERCAVFCQACQACHAVLADS 90
DB 49 ASMWSSMSRSDASV---ADMERAQL--EBEAAAAB-----ERPLVFLCSCGRRPLGDS 95
QY 91 VHLAWDLRS-SLGAVVFSRVTVNNVLEAPFLVGIEGSLKSTYNNLLFCGSGCIPVGFHLY 149
DB 96 --LSWASQEDTNCILLRCVSCNVSDKEQKLSKREKNGCVLETLCCAGCSLNLGYVYR 153
QY 150 STHAALALRGHFLCSSLDDKMWCVYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
DB 154 CTPKNLDYKRDLPCLSVBAIESYVLGSSEKQIV-SEDKELFNL---ESRVEIKSLTQME 209
QY 208 NRLKSLMKILSE 219
DB 210 DVLKALQMKLWE 221

RESULT 7
ADJ69163
ID ADJ69163 standard; protein; 233 AA.
XX AC ADJ69163;
XX DT 06-MAY-2004 (first entry)
XX DE Human heat mitochondrial protein as a therapeutic target SeqID969.
XX KW mitochondrial; human; screening assay; diabetes mellitus;
XX KW Huntington's disease; osteoarthritis;
XX KW Leber's hereditary optic neuropathy; LHON;
XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX KW neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
XX KW osteopathic; ophthalmological; cytoskeletal.
XX OS Homo sapiens.
XX PN WO2003087768-A2.
XX PD 23-OCT-2003.
XX PF 04-APR-2003; 2003WO-US010870.
XX PR 12-APR-2002; 2002US-0372843P.
XX PR 17-JUN-2002; 2002US-0389987P.
XX PR 20-SEP-2002; 2002US-0412418P.
XX PA (MITO-) MITOKOR.
XX PA (BUCK-) BUCK INST AGE RES.
XX PA Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
XX PI Warnock DE;
XX PR WPI; 2003-845369/78.
XX DR Identifying a mitochondrial target for drug screening assays and for
XX PT treating diseases associated with altered mitochondrial function,
XX PT comprises detecting a modified polypeptide in a sample and correlating
XX PT with the disease.
XX PS Claim 1; SEQ ID NO 969; 180pp; English.
XX CC This invention relates to novel mitochondrial targets that can be used
XX CC for therapeutic intervention in treating a disease associated with
XX CC altered mitochondrial function. Specifically, it refers to a method for
XX CC identifying proteins of the human heart mitochondrial proteome that are
XX CC useful for drug screening assays, as well as therapeutic targets. The

CC present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, neurotropic, antidiabetic, anticonvulsant, antiarthritic, osteopathic, ophthalmological and cytoskeletal activities. This polypeptide sequence is a human heart mitochondrial protein of the invention.

Sequence 233 AA;

Query Match 9.4%; Score 113; DB 7; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0011;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTTSMEDTQVVGSSPLGAGLGAEPAPAGPOLPSWLQPERCAVFCQACQACHAVLADS 90
DB 49 ASMWSSMSRSDASV---ADMERAQL--EBEAAAAB-----ERPLVFLCSCGRRPLGDS 95
QY 91 VHLAWDLRS-SLGAVVFSRVTVNNVLEAPFLVGIEGSLKSTYNNLLFCGSGCIPVGFHLY 149
DB 96 --LSWASQEDTNCILLRCVSCNVSDKEQKLSKREKNGCVLETLCCAGCSLNLGYVYR 153
QY 150 STHAALALRGHFLCSSLDDKMWCVYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
DB 154 CTPKNLDYKRDLPCLSVBAIESYVLGSSEKQIV-SEDKELFNL---ESRVEIKSLTQME 209
QY 208 NRLKSLMKILSE 219
DB 210 DVLKALQMKLWE 221

RESULT 8
ADJ31517
ID ADJ31517 standard; protein; 233 AA.
XX AC ADJ31517;
XX DT 20-MAY-2004 (first entry)
XX DE Human protein encoded by a full length cDNA clone SeqID 3550.
XX KW human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method.
XX OS Homo sapiens.
XX PN EP1396543-A2.
XX PD 10-MAR-2004.
XX PF 07-JUL-2000; 2003EP-00025638.
XX PR 08-JUL-1999; 99JP-0019486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183665.
XX PR 07-JUL-2000; 2000EP-00114089.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX PR WPI; 2004-204755/20.
XX DR N-PSDB; ADJ31516.
XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
XX PS length human cDNAs.
XX CC Example 1; SEQ ID NO 3550; 1340pp; English.

CC This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction.
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.

XX SQ Sequence 233 AA;

Query Match 9.4%; Score 113; DB 8; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0011;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
QY 31 ASFTTSEWDFTQVVGKSPGLGAEPAAGPQLPSWLQPERCAVFOCAQCHAVLADS 90
DB 49 ASMWSSMEDASV-----ADMERAGL--EEEAARAE-----ERPLVLCGCRPLGDS 95
QY 91 VHLAWDLGR-SLGAVFVSRTNNVLEAPFLVIGIEGLSGTYNLLFCGSGCIPVGFHLY 149
DB 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKENGCVLETLCCAGCSLNLGYVYR 153
QY 150 STHAALAAALRGHCLSSDKMVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
DB 154 CTPKRLDYKRDLCFLSVEATESYVIGSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209
QY 208 NRLKSLMKILSE 219
DB 210 DVLKALQKLWE 221

RESULT 9

ABBS4167
ID ABB54167 standard; protein; 695 AA.

XX AC ABB54167;
XX DT 29-AUG-2003 (revised)
XX DT 16-MAY-2002 (first entry)
XX DE Lactococcus lactis protein yihC.
XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX OS Lactococcus lactis; IL1403.

XX PN FR2807446-A1.

XX PD 12-OCT-2001.

XX PF 11-APR-2000; 2000FR-00004630.

XX PR 11-APR-2000; 2000FR-00004630.

XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX DR WPI; 2002-043418/06.

XX PT New nucleotide sequence useful in the identification or Lactococcus
XX PT lactis and related species.

XX PS Claim 6; SEQ ID NO 869; 2504pp; French.

XX CC The present invention is related to a Lactococcus lactis nucleotide
XX CC sequence (AB90521) and related proteins (ABB5300-ABB55621). The nucleic
XX CC acid sequence is useful in the detection and/or amplification of nucleic
XX CC acid sequence, particularly to identify Lactococcus lactis or related
XX CC species. The proteins of the invention are useful for the biosynthesis or
XX CC biodegradation of a composition of interest. The invention helps research

CC in lactic bacteria, particularly useful in the production of yogurt and
CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)

XX SQ Sequence 695 AA;

Query Match 7.7%; Score 92; DB 5; Length 695;
Best Local Similarity 25.3%; Pred. No. 1.5;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
QY 101 LGAVFVSRTNNVLEAPFLVIGIEGLSGTYNLLFCGSGCIPVGFHLYSTHAALALRG 160
DB 101 LGTIIF-----FVSGTFFFGAGELSKRKPAMMLITMGITVAY-AYSVIATIMSLNG 153
QY 161 HF-----CLSSDKMVCYLLKTKAIVNASE--MDI-----QNVPLSE-K 195
DB 154 HMGNFVPELATLIVIMLIGHLEMKAIMGAGDALDASLVPKKAHLKSGKDVSELK 213
QY 196 IABELKEKIVLTHNRLKSLMKILSEVTPDQS 225
DB 214 VGDL--LLVKEKEKIPADGLILSEALVDES 241

RESULT 10

ADS29356
ID ADS29356 standard; protein; 695 AA.

XX AC ADS29356;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #18389.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 18389; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a

PI Legace RE, Spiro BA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patury S, Shi X, Suarez CJ;
 XX WPI; 2004-329368/30.
 DR N-PSDB; ACN42006.
 XX
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 PS Claim 27; Page; 190pp; English.
 XX
 XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 572 AA;

Query Match 7.4%; Score 89; DB 8; Length 572;
 Best Local Similarity 23.0%; Pred. No. 2.4;
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
 QY 8 HRSRCATPPRGDFC-----GATERAIDQ-----SFTSMEDTQVVGSSPL----- 50
 DB 342 YRQQAQPPH---CPAPEGEPGAPQALGDAPTSVSLTTAVQDVVCPQSGSHALCTCCFQ 398
 QY 51 ---GPAGLGAEEPAAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
 DB 399 PMPDRAEREQDPRVAPQ-----QCAVC---LQPFCHLYWGCTRTGCGYCLA 442
 QY 100 -----SLGAVVFSRVNNVLEAPFLVIGEGSLKGSTY-NLLFCGSGGIPVGFHLVSTHA 153
 DB 443 PFCELNLGDKCLDGLVNNNSYESDILKNYLAT-RGLTWKML-----TES 486
 QY 154 ALAALRGHFCLS-----SDKMVCYLLKTKAIVNASEMDIONVPLSE 194
 DB 487 LVALQRGVFLLSDYRVVTGDTVLCYCGLRSFRELITYQOONIPASE 532

RESULT 13
 AAB93182
 ID AAB93182 standard; protein; 623 AA.
 XX
 XX AAB93182;
 XX
 XX 26-JUN-2001 (first entry)
 XX
 XX Human protein sequence SEQ ID NO:12128.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 XX Homo sapiens.
 XX
 XX EP1074617-A2.
 XX
 XX 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12128; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

Sequence 623 AA;

Query Match 7.4%; Score 89; DB 4; Length 623;
 Best Local Similarity 23.0%; Pred. No. 2.7;
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GATERAIDQ-----SFTSMEDTQVVGSSPL----- 50
 DB 393 YRQQAQPPH---CPAPEGEPGAPQALGDAPTSVSLTTAVQDVVCPQSGSHALCTCCFQ 449
 QY 51 ---GPAGLGAEEPAAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
 DB 450 PMPDRAEREQDPRVAPQ-----QCAVC---LQPFCHLYWGCTRTGCGYCLA 493
 QY 100 -----SLGAVVFSRVNNVLEAPFLVIGEGSLKGSTY-NLLFCGSGGIPVGFHLVSTHA 153
 DB 494 PFCELNLGDKCLDGLVNNNSYESDILKNYLAT-RGLTWKML-----TES 537
 QY 154 ALAALRGHFCLS-----SDKMVCYLLKTKAIVNASEMDIONVPLSE 194
 DB 538 LVALQRGVFLLSDYRVVTGDTVLCYCGLRSFRELITYQOONIPASE 583

RESULT 14
 ABB97233
 ID ABB97233 standard; protein; 623 AA.
 XX
 XX ABB97233;
 XX


```
DT 27-JUN-2002 (first entry)
XX
XX
DE Novel human protein SEQ ID NO: 501.
XX
XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
XX antiinfertility; cerebroprotective; cytosatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX Homo sapiens.
XX
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
XX
XX N-PSDB; ABN32419.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Example 2; SEQ ID NO 501; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention
XX
XX Sequence 623 AA;
XX
XX Query Match 7.4%; Score 89; DB 5; Length 623;
XX Best Local Similarity 23.0%; Pred. No. 2.7;
XX Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
XX
XX QY 8 HRSRCATPPRGDFC-----GGTERAIDQA-----SFTTSMEDWTQVVKGSPL----- 50
XX Db 393 YRQAAQPPH---CPAPEGEPGAPQALGDPPTSLSLTAVQDYVCPLOGSHALCTCCQ 449
XX
XX QY 51 ---GPAGLGAEEPAAGPQLPSMLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
XX Db 450 PMPDRAEREQDPVAPQ-----QCAVC---LQPFCHLYWCCTRTGCGYGLA 493
XX
XX QY 100 -----SLGAVVFSRVNTNNVLEAPFLVIGESLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
XX Db 494 PFCELNLGDKCLDGVLLNNNSYESDILKNYLAT-RGLTWKNML-----TES 537
XX
XX QY 154 ALAALRGHFCLS-----SPKWCYLLKTKAIVNASSEMDIQNVPLSE 194
XX Db 538 LVVALQGVFLLSDYRVGTGTVLCYCCLGRSFRLETYYQOONIPASE 593
XX
XX RESULT 15
XX AAB93168
XX ID AAB93168 standard; protein; 652 AA.
XX
XX AAB93168;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12100.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
XX
XX 11-JAN-2000; 2000JP-00118776.
XX
XX 02-MAY-2000; 2000JP-00183767.
XX
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesising polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 652 AA;
XX
XX Query Match 7.4%; Score 89; DB 4; Length 652;
XX Best Local Similarity 23.0%; Pred. No. 2.9;
XX Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
XX
XX QY 8 HRSRCATPPRGDFC-----GGTERAIDQA-----SFTTSMEDWTQVVKGSPL----- 50
XX Db 422 YRQAAQPPH---CPAPEGEPGAPQALGDPPTSLSLTAVQDYVCPLOGSHALCTCCQ 478
XX
XX QY 51 ---GPAGLGAEEPAAGPQLPSMLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
XX Db 479 PMPDRAEREQDPVAPQ-----QCAVC---LQPFCHLYWCCTRTGCGYGLA 522
XX
XX QY 100 -----SLGAVVFSRVNTNNVLEAPFLVIGESLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
XX Db 523 PFCELNLGDKCLDGVLLNNNSYESDILKNYLAT-RGLTWKNML-----TES 566
```

Oy 154 ALAALRGHFCLS-----SDXWVCYLLKTKAIVNASEMDIQNVPLSE 194
Db 567 LVALORGVELLSDYRVGTDLVCYCCGLRSFRELTYQYRQNPASE 612

Search completed: March 25, 2006, 07:59:14
Job time : 191 secs

A:Molecule type: mRNA
A:Residues: 1-2194 <GRE>
A:Cross-references: UNIPROT:Q03460; UNIPARC:UPI000012B7FA; GB:L01660; NID:gl166411; PIDN:
A:Accession: PQ0551
A:Molecule type: protein
A:Residues: 102-114 <GR2>
A:Cross-references: UNIPARC:UPI0000172061
C:Comment: This enzyme catalyzes the reductive transfer of the amido group of glutamine
C:Superfamily: glutamate synthase (NADH/NADPH), eukaryotic type
C:Keywords: 3Fe-4S; chloroplast; iron-sulfur protein; metalloprotein; NAD; oxidoreductase
F:1-101/Domain: propeptide #status predicted <PRO>
F:102-2194/Product: glutamate synthase (NADH) #status experimental <MAT>
F:102/Active site: Cys #status predicted
F:1246,1252,1257/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 7.1%; Score 85; DB 1; Length 2194;
Best Local Similarity 24.7%; Pred. No. 39;
Matches 58; Conservative 22; Mismatches 79; Indels 76; Gaps 12;
QY 4 OPLRHRSRCATPPRGDFCGGTERAIDQAS-----FTTSMEDTQVVKSSPLGPA 53
DB 1037 EPLADCSR--NPKRS-----AIKQVAGRGFGVSSVYLTNADELQIKMAQAGAKP----- 1082

QY 54 GLGABEP-----AAGPQFSLQPERCAVFCQAQCHAVLADSVHLAWDLRS 99
DB 1083 GEGGELPGHKVIGDIAITRNSTAGVGLIS--PPPHDIYS-----IEDLAQLIHLNKN 1133

QY 100 SLGAVFSRVNNVLEAPFLVIGIEGSLKSTYNLLFCG-----SCGIPVG 145
DB 1134 ANPA---ARISVKLVSEAGVGVASGVVKGHAHVLSIGHDGGTGASRWGTGKSGAGLPWE 1190

QY 146 FHLVSTHAALAA--LRGHFCLSSDKMVCYLLKT-----KAIVNASEMDIQNVPL 192
DB 1191 LGLAETHQTLVANDLRGRHTTLQDQ-----LKTGRDVVAIALAGAEYGFSTAPL 1241

RESULT 7
AE2074
ferrichrome binding protein of ABC transporter all2147 [imported] - Nostoc sp. (strain F
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE2074
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2074
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <KUR>
A:Cross-references: UNIPROT:O8VV34; UNIPARC:UPI000000CE32E; GB:BA000019; PIDN:BA073846.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2147
C:Superfamily: ferrichrome-iron transport protein feCB

Query Match 7.1%; Score 84.5; DB 2; Length 361;
Best Local Similarity 26.1%; Pred. No. 4.3;
Matches 36; Conservative 17; Mismatches 36; Indels 49; Gaps 7;
QY 97 LRSRLGAVVF---SRVT---NNVLEAPFLVIGIEGSLKSTYNLLFCGSC-----GIPV 144
DB 81 ISHALGKVKIPLKQPVVVLEENILDSVLALGVK-----PVGMYQCQCEENFRGIP- 133

QY 145 GFHLYSTHAALALRGHFLCSSDKMVCYLLKTKAIVNASEMDIQNVPLSEKIAELKEKIV 204
DB 134 -----SD-----LLADVPVVG-----NIGNQPSLEKILSLKPDLI 163

QY 205 LTHNRLKSLMKILSEVTP 222
DB 164 LGTLWLSKSYKILSSIAAP 181

RESULT 8
D97047
DNA ligase (NAD dependent), Liga [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004
C:Accession: D97047
R:Noelling, J.; Brenon, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Le
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-663 <KUR>
A:Cross-references: UNIPROT:O97J58; UNIPARC:UPI000000CA10F; GB:AE001437; PIDN:AAK79167.
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1195
C:Superfamily: DNA ligase (NAD), Liga type

Query Match 6.9%; Score 83; DB 2; Length 663;
Best Local Similarity 28.6%; Pred. No. 13;
Matches 36; Conservative 18; Mismatches 44; Indels 28; Gaps 7;
QY 102 GAVVFSRVNNVLEAPFLVIGIEGSLKSTYNLL-----LFCGSCGIPVGFHLYSTHAAL 155
DB 368 GARVFLRRSNDVI---PEIMGVTEETEGTEKEIRAPTICPYCGSBIKGVGHL----- 417

QY 156 AALRGHFLCSSDKMVCYLLKTKAIVNASEMDIQNVPL--LSEKIAE--LKEKIVLTHNRLKSL 213
DB 418 -----FC--ENTLSCKPQWYKSVHVFASRKAMNIEGFSEKTAQLFEK--LNKISISDL 467

QY 214 MKILSE 219
DB 468 YRIKSE 473

RESULT 9
B69099
formate hydrogenlyase, iron-sulfur subunit 2 - Methanobacterium thermoautotrophicum (st
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69099
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, I
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fun
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69099
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-143 <MTH>
A:Cross-references: UNIPROT:O27769; UNIPARC:UPI00000066781; GB:AE0000929; GB:AE0000666; NJ
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1736
C:Superfamily: nrfC protein; ferredoxin 2 [4Fe-4S] homology
F:62-118/Domain: ferredoxin 2 [4Fe-4S] homology <FER2>

Query Match 6.9%; Score 82.5; DB 2; Length 143;
Best Local Similarity 23.0%; Pred. No. 2.1;
Matches 35; Conservative 26; Mismatches 66; Indels 25; Gaps 6;
QY 72 PERCAVFCQAQCHAVLADSVHLAWDLRSIGAVVF-----SRVTNNVLEAPFLVIGIEG 125
DB 8 PELCD--ECKMCRICPKNAIRVID-----GVFVCMHCSPERPAPCLNIFEDAIIVSDG 60

QY 126 SLKGSTYNLLFCGSC--GIPVGFHLYSTHAALALRGHFLCSSDKMVCYLLKTKAIVNAS 183

Db 61 AVVILEDRICGCLCRDACPVG--AITLNERGVAVKDLICIDRDKPLCMVCPKGALES 118
Qy 184 EMDIQNVPLSEKIAELKKEKIVLTHNRLKSLMK 215
Db 119 SEDM-----MAAKRDKIAGELKRLKMLK 142

RESULT 10
A5433
sugar transporter like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A85433
R:anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: A85433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <STO>
A:Cross-references: UNIPROT:O23213; UNIPARC:UPI000009F22F; GB:NC_001268; NID:G7270615; E
C:Genetics:
A:Map position: 4
A:Superfamily: glucose transport protein

Query Match 6.9%; Score 82.5; DB 2; Length 493;
Best Local Similarity 24.9%; Pred. No. 10;
Matches 53; Conservative 32; Mismatches 59; Indels 69; Gaps 13;

Qy 79 QCAQCHAVLADSVHLAWLSRLGAVVP---SRVTNNVLEAPFLVGI-----EGSLK- 128
Db 19 QCAIVASIV--SIIPGYDTGVMSGAMVFIEDLTKTDQIEV--LTGTLNLCALVGSLLA 74

Qy 129 -----GSTYNLLFCGSC--GIPVGFHL-----YSTH 152
Db 75 GRTSDIIGRRYTVIVLASILFMLGSLMGWPNYPVLLSGRTAGLVGFGALMVAPVSAE 134

Qy 153 AALAALRG-----HFCLSSDRWCYLLK---TKAIVNAS---EMDIQNVLP---LSEKIA 197
Db 135 IATASHRGLLASPLHLCISIGILLGYIVNYPFSLKPMHIGWRLMGLIAAAPSILVAFGIL 194

Qy 198 ELKE--KIVLTHNRLKSLMKILSEV--TPDQSK 226
Db 195 KMPSPRWLIMQGRLEKGEKILELVSNSPESA 227

RESULT 11
F87678
DNA polymerase I [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: F87678
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Koon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87678
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-967 <STO>
A:Cross-references: UNIPROT:Q9A2U2; UNIPARC:UPI00000C7AB1; GB:AE005673; NID:gi3425184; E
C:Genetics:
A:Gene: CC3464
C:Superfamily: DNA-directed DNA polymerase I

Query Match 6.8%; Score 82; DB 2; Length 967;
Best Local Similarity 24.1%; Pred. No. 26;
Matches 51; Conservative 27; Mismatches 72; Indels 62; Gaps 11;

Qy 12 CATPRGDFCGGTERAIDQ---ASFTTSMWDVT---QVVKSSPLGPRAGLGAEBEPAAGPQ 65
Db 271 CDTPLPQPLDALTVREPDKALAAFLQMEFRSLARRVGDGSAATPGTL--DRPAAPPK 328

Qy 66 LP-----SWL-----OPERC--AVFOCAQCHAVLADSVHLAWLSRLGAVVFSR 108
Db 329 APVSVSYNGAAARAAHPVPVKIDHAYACVRDLATLKAWAKATD-----KGLVAFDT 384

Qy 109 VTNNVLEAPFLVGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAAALRGHFLCSSDK 168
Db 385 ETD-----ALSSAT-----AGLCGV-----SLAIAPCEACVIP-- 412

Qy 169 MCVYLLKTKAIVNASEMDIQNVPLSEKIAELK 200
Db 413 -ISHCEKADGLAFAPADIEQIPLADVIATLK 443

RESULT 12
D90404
transport protein, probable [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: D90404
R:She, O.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Char
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-NGOC, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: D90404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <KUR>
A:Cross-references: UNIPROT:Q97W97; UNIPARC:UPI00000646FA; GB:AE006641; NID:gl3815639;
C:Genetics:
A:Gene: SS02338

Query Match 6.8%; Score 81.5; DB 2; Length 329;
Best Local Similarity 22.2%; Pred. No. 7.4;
Matches 40; Conservative 30; Mismatches 47; Indels 63; Gaps 7;

Qy 87 LADSVHLA-WD-----LSRSLGAVVFSRVTNNV-----LE 116
Db 33 LSESHLAYMEVFAIVLALPFLGRIIGSFIVQFKNVSYCFPLGLVILQNLGALIF 92

Qy 117 APFLVG-----IEGSLKSTYNLLFCGSCGIPVGFHLYSTHAAALRGHFLCSS 166
Db 93 VRFLVGVIIFGLLTSYAVESAVKSGRNVLVGTTAGWPIGWVI----- 134

Qy 167 DKWVCY-LLKTKAIVNASEMDIQNVPLSE---KIAELKEKIVLTHNRLKSLMKILSEVTP 222
Db 135 -SYVAYVLLKNWNVINISGILMLLALFELNGKEFGERSKISVSFPRLTSLIYVSALT 193

RESULT 13
H85767
Glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H85767
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouais, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85767
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: UNIPROT:Q8X673; UNIPARC:UPI00000DDBE5; GB:AE005174; NID:gi2515601;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: uidB

C;Superfamily: melibiose carrier protein

Query Match 6.8%; Score 81.5; DB 2; Length 457;

Best Local Similarity 25.2%; Pred. No. 11;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIAASIGQGVMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374
QY 142 IPVGPHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGLSGYIANQVTPVIMGIRTSIALVPCGFMLLAFVIWFFPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221

Db 432 K--KPKKEIVVEIDNRKKVQQQLISDIT 456

RESULT 14

C90919

glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain RMD 05099

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: C90919

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C90919

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-457 <HAY>

A;Cross-references: UNIPROT:Q8X673; UNIPARC:UPI00000D0B85; GB:BA000007; PIDN:BAB35746.1;

A;Experimental source: strain O157:H7, substrain RMD 0509952

C;Genetics:

A;Gene: EC52323

C;Superfamily: melibiose carrier protein

Query Match 6.8%; Score 81.5; DB 2; Length 457;

Best Local Similarity 25.2%; Pred. No. 11;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIAASIGQGVMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGPHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194

Db 375 QAIG--GSIPAFILGLSGYIANQVTPVIMGIRTSIALVPCGFMLLAFVIWFFPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221

Db 432 K--KPKKEIVVEIDNRKKVQQQLISDIT 456

RESULT 15

B64918

glucuronide permease uidB - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: B64918

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64918

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-457 <BLAT>

A;Cross-references: UNIPROT:P30868; UNIPARC:UPI00000137B23; GB:AE000257; GB:U00096; NID:9

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: uidB; gusB

C;Superfamily: melibiose carrier protein

F;37-53/Domain: transmembrane #status predicted <TM01>

F;81-97/Domain: transmembrane #status predicted <TM02>

F;152-168/Domain: transmembrane #status predicted <TM03>

F;184-200/Domain: transmembrane #status predicted <TM04>

F;231-247/Domain: transmembrane #status predicted <TM05>

F;263-279/Domain: transmembrane #status predicted <TM06>

F;310-326/Domain: transmembrane #status predicted <TM07>

F;408-424/Domain: transmembrane #status predicted <TM08>

Query Match

Best Local Similarity 6.8%; Score 81.5; DB 2; Length 457;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIAASIGQGVMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGPHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194

Db 375 QAIG--GSIPAFILGLSGYIANQVTPVIMGIRTSIALVPCGFMLLAFVIWFFPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221

Db 432 K--KPKKEIVVEIDNRKKVQQQLISDIT 456

Search completed: March 25, 2006, 08:03:57.

Job time : 42 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 07:56:15 ; Search time 231 Seconds
(without alignments)
699.420 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAAPQLRHRSRCATPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	1 OIP5 HUMAN	O43482 homo sapien
2	122	10.2	218	2 O9CXK6 MOUSE	O9CXK6 mus musculus
3	118	9.8	462	2 O4PAB5 USTMA	O4PAB5 ustilago ma
4	115.5	9.6	204	2 O9CZJ6 MOUSE	O9CZJ6 mus musculus
5	113	9.4	233	1 CU045 HUMAN	O9NYP9 homo sapien
6	113	9.4	233	2 O54220 HUMAN	O54220 homo sapien
7	110.5	9.2	232	1 CU045 PANTR	O68ut5 pan troglod
8	109	9.1	207	2 O5BLBI BRARE	O5BLBI brachydanio
9	107	8.9	532	2 O54HI6 DICDI	O54HI6 dictyoscell
10	105	8.8	155	1 YCZC SCHPO	O9P802 schizosacch
11	94	7.8	1556	2 O4Q626 LEIMA	O4Q626 leishmania
12	92	7.7	361	2 O8S611 ORYGA	O8S611 oryza sativ
13	92	7.7	695	2 O9CH87 LACLA	O9CH87 lactococcus
14	91	7.6	628	2 O69KE0 ORYSA	O69KE0 oryza sativ
15	90.5	7.6	597	2 O83K93 TREPA	O83K93 treponema p
16	89.5	7.5	502	2 O50PD2 ENTHI	O50PD2 entamoeba h
17	89.5	7.5	661	2 O50V43 ENTHI	O50V43 entamoeba h
18	89.5	7.5	661	2 O50WZ8 ENTHI	O50WZ8 entamoeba h
19	89	7.4	633	2 O5M052 STRT1	O5M052 streptococc
20	89	7.4	664	1 CHER HUMAN	O96ep1 homo sapien
21	89	7.4	735	2 O5M4R3 STRY2	O5M4R3 streptococc
22	89	7.4	893	1 POL2 BAMWA	O65329 barley mild
23	88.5	7.4	152	2 O50XW7 ENTHI	O50XW7 entamoeba h
24	88.5	7.4	638	2 O75RZ6 CIOIN	O75RZ6 ciona intes
25	88	7.3	250	1 KKA7 CRAWJE	P14508 campylobact
26	88	7.3	493	2 O68VH2 9POTY	O68VH2 barley mild
27	87.5	7.3	634	2 O6CX82 KLUJA	O6CX82 kluyveromyc
28	87.5	7.3	636	2 O4UTS4 CORUK	O4UTS4 corynebacte
29	87	7.3	664	1 CHER MOUSE	O81013 mus musculus
30	86.5	7.2	669	2 O97FQ5 CLOAB	O97FQ5 clostridium
31	86.5	7.2	2093	2 O5K2K3_9RHAB	O5K2K3 isfahan vir

ALIGNMENTS

RESULT 1

OIP5_HUMAN
ID OIP5_HUMAN STANDARD; PRT; 229 AA.
AC O43482; Q96BX7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Opa-interacting protein 5.
GN Name=OIP5;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98125741; PubMed=9466265;
RA Williams J.M., Chen G.-C., Zhu L., Rest R.F.;
RT "Using the yeast two-hybrid system to identify human epithelial cell
RT proteins that bind gonococcal Opa proteins: intracellular gonococci
RT bind pyruvate kinase via their Opa proteins and require host pyruvate
RT for growth.";
RL Mol. Microbiol. 27:171-186(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Rosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttzfeld Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: Binds outer membrane protein Opa from Neisseria
CC gonorrhoeae.
CC -!- INTERACTION:
CC P04049:RAFI; NBExp=3; IntAct=EBI-536879, EBI-365996;
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

32 86.5 7.2 5344 2 Q529N3 MAGGR Q529N3 magnaportha
33 86 7.2 145 2 Q839T4 ENTFA Q839T4 enterococcu
34 86 7.2 327 2 Q4NE64 9MICC Q4NE64 arthrobacte
35 86 7.2 331 2 Q6GXE5 COCCA Q6GXE5 cochllobolu
36 86 7.2 893 2 Q68VH4 9POTY Q68VH4 barley mild
37 86 7.2 3692 2 Q5J1O7_9NOCA Q5J1O7 nocardia un
38 85 7.1 337 1 SW1 STRAW Q82hul streptomycy
39 85 7.1 671 2 Q815K5 PLAF7 Q815K5 plasmodium
40 85 7.1 1018 2 Q6FQK6 CANGA Q6FQK6 candida gla
41 85 7.1 2194 1 GLSN MEDSA Q03460 medicago sa
42 85 7.1 2194 2 Q40360 MEDSA Q40360 medicago sa
43 84.5 7.1 361 2 Q8YV34 ANASP Q8YV34 anabaena sp
44 84.5 7.1 376 2 Q5YZE1 NOCFA Q5YZE1 nocardia fa
45 84 7.0 247 2 Q64WS0 BACFR Q64WS0 bacteroides

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; AF025441; AAC39561.1; ALT INIT; mRNA.
 DR EMBL; BC015050; AAHL5050.1; -; mRNA.
 DR Inact; 043482; -; mRNA.
 DR Ensembl; ENSG00000104147; Homo sapiens.
 DR HGNC; HGNC:20300; OIP5.
 DR H-InvDB; HIX0012150; -.
 DR MIA; 606020; -.
 DR GO; GO:0005515; F:protein binding; TAS.
 DR GO; GO:0007154; P:cell communication; NAS.
 DR GO; GO:0007154; P:cell communication; NAS.
 SQ SEQUENCE 229 AA; 24691 MW; 0EBD4006193A3106 CRC64;

Query Match 100.0%; Score 1198; DB 1; Length 229;
 Best Local Similarity 100.0%; Pred. No. 8.8e-105;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAAGPLHRRCATPPRCDCGCTERAIDQASFTTSNEWDTQVVGSSPLGAGABEP 60
 DB 1 MAAGPLHRRCATPPRCDCGCTERAIDQASFTTSNEWDTQVVGSSPLGAGABEP 60
 QY 61 AAGPQLPSWLOPERCAVFCQAQCHAVLADSVHLAWLDSRLGAVFGRVNNVLEAPFL 120
 DB 61 AAGPQLPSWLOPERCAVFCQAQCHAVLADSVHLAWLDSRLGAVFGRVNNVLEAPFL 120
 QY 121 VGIEGSLKSTYNLLFCGSCGIPVGFHLYTHAALALRGHFLCLSSDKWCVYLLKTKAIV 180
 DB 121 VGIEGSLKSTYNLLFCGSCGIPVGFHLYTHAALALRGHFLCLSSDKWCVYLLKTKAIV 180
 QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQSKPEN 229
 DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQSKPEN 229

RESULT 2

Q9CXR6 MOUSE
 ID Q9CXR6_MOUSE PRELIMINARY; PRT; 218 AA.
 AC Q9CXR6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched
 DE library, clone:3110025H23 product:hypothetical protein, full insert
 DE sequence.
 GN Name=2610039C10Rik;
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P.; Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh A.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamanaka R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa Y., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK014084; BAB29147.1; -; mRNA.
 DR Ensembl; ENSMUSG00000022978; Mus musculus.
 DR MGI; MGI:1913828; 2610039C10Rik.
 KW Hypothetical protein.
 SQ SEQUENCE 218 AA; 24447 MW; 5539085BFDDDE8C2 CRC64;

Query Match 10.2%; Score 122; DB 2; Length 218;
 Best Local Similarity 23.3%; Pred. No. 0.0049;
 Matches 48; Conservative 30; Mismatches 102; Indels 26; Gaps 6;
 QY 29 DQASFTTSMEDTQVVGSSPLGAGABEPAGPQLPSWLOPERCAVFCQAQCHAVLA 88
 DB 15 DSSRYLRLOKWAN--MSSDALGLVLEKPERPEKAAANP-----LVFLCARCRPLG 64
 QY 89 DSVHLAWLDSR--SLGAVFGRVNNVLEAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFH 147
 DB 65 DS--LTWVASQEDTNCILLRSVCSNVSDKEPKLSKCRDDEGCLLEALYCTGCSLSGVY 122

QY 148 LYSTHAALALRGHPCILSSDKWCVLL-----KTKAIVNASEMDIQNVPLSEKTA 198
 Db 123 YRCTPKNDYKRDULFCLSVETVSYLGSSEKQIVSKDKELFNLE---SRVEIEKSIQ 178
 QY 199 LKEKIVLTHNRLKSLMKILSEVTPDQ 224
 Db 179 MEEVLTALQKLRVESKLSLAQPCQ 204

RESULT 3

Q4PAB5 USTWA
 ID Q4PAB5 USTWA PRELIMINARY; PRT; 462 AA.
 AC Q4PAB5
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=UM02948.1;
 OS Ustilago maydis 521.
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=237631;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=521;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
 RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
 RA Borowsky M., Boukhaltel B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David K., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gierre S.,
 RA Girke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heiler A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huby E., Iliev I.,
 RA Jaffe D., Jones C., Kamat M., Kamat A., Kamvesselis M., Karlsson E.,
 RA Kells C., Kieu A., Kigner P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., Maru K., Matthews C., Mauceli E.,
 RA Manning J., Marbella R., McConough S., McGhee T., Meldrum J., Menus L.,
 RA McCarthy M., Mcdonough S., McGhee T., Meldrum J., Menus L.,
 RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotoho B.,
 RA O'neil K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
 RA Tenzing P., Testaye S., Theodore J., Thoulutsang Y., Topham K.,
 RA Towey S., Teamlu T., Tsomo N., Vallee D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Zander E.;

RT "The genome sequence of Ustilago maydis.";
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC ENBL; AACP01000101; EAK84120.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 462 AA; 49283 MW; EDDCFDF17757B509 CRC64;

Query Match 9.8%; Score 118; DB 2; Length 462;
 Beat Local Similarity 25.9%; Pred. NO. 0.03;
 Matches 64; Conservative 28; Mismatches 109; Indels 46; Gaps 10;
 QY 2 AAQPLRHRSRCATPRGDFCGTERTAIQAQSFSTTSMENDTQV-VKGSSPLGAGLGAEP 60
 Db 109 AAQPLSSDDDDMAFPKG-----GSLTTT---STRARGRGGRGPRGRGRG 153
 QY 61 AA-----GPQL--PSWLOPERCAVFCQACHAVLADSV-HLAWDLSSRLGAVVFSRYT 110
 Db 154 ASTSTTVVHKPDLESSESEPPNPLVFCQRCCLGSLAFVATDI--DLGVILSDVS 211
 QY 111 NNVLREAPPLVGIE-GSLKSGTYNLLFCGSGCIPVGFHYLSTHAALALRGHFCILSSDQM 169
 Db 212 EIIQDDTTYETSTETPGKDIGSTFARLCAGCAAAVGRNYRTTPRDLDLRDCFSLEVD 271
 QY 170 VCYLLKTKAIVNASEMDIQ-----NVLSEKIAELKEKIVLTH-NRLKS 212
 Db 272 YTYQSGSNYTRQEEDEDDQDALVDGASGKPARASRPTHETVQAEKIVLTKMERT 331
 QY 213 LMKILSE 219
 Db 332 LTIELSD 338

RESULT 4

Q9CZJ6 MOUSE
 ID Q9CZJ6 MOUSE PRELIMINARY; PRT; 204 AA.
 AC Q9CZJ6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Mus musculus 11 days embryo whole body cDNA, RIKEN full-length
 DE enriched library, clone:2700078124 product:hypothetical protein, full
 DE insert sequence (2610039C10Rik protein).
 GN Names=2610039C10Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RN Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Azakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schramm L.M., Stauberli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body; DOI=10.1038/nature01266;
RX MEDLINE=22354683; PubMed=12466851; PubMed=12466851; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakaio I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Mateuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia K., Corbani L.E., Cousins S.,
RA Della E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaestland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochin I.V., Lee Y., Lemhard B., Lyons P.A.,
RA Meglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arahata T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630 (2000).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771 (2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Akawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijina Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
[7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Eye;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
[8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Eye;
RA Director MGC Project;
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK012533; BAE28302.1; -; mRNA.
DR EMBL; BC079900; AAH79900.1; -; mRNA.
DR Ensembl; ENSMUSG0000022978; Mus musculus.
DR MGI; MGI:1913828; 261003C10Rik.
KW Hypothetical protein.
SQ SEQUENCE 204 AA; 22949 MW; BBD667250FDF9B0F CRC64;

Query Match 9.6%; Score 115.5; DB 2; Length 204;
Best Local Similarity 25.6%; Pred. No. 0.018;
Matches 50; Conservative 29; Mismatches 97; Indels 19; Gaps 7;

QY 29 DQASFTTSMENDTVQVKGSGPLGPAEPAAGPQPSWLOPERCAVQCACQAVLA 88
DB 15 DSSRYLRLOKQW--MSSADALGLEKPERPEKAAANP-----LVFLCARCRPLG 64
QY 89 DSVHLAWDLRS-SLGAVVFGRVTNNVLEAPFLVGIEGSLKSTYNTLLFCGSCGIPVGFH 147
DB 65 DS--LTFVASQEDTNCILLRSVCSVDKPEKLSKCRDEGCGILEALYTCGSLSGYV 122
QY 148 LYSTHAALAAALRGHGLSSDQWVCYLL--KTKAIVNASEMDIONVPLSEKIAELKEKIVL 205
DB 123 YRCTPKNLDYKRDLCFLSVEAVESYTLGSGEKIV-SEDKELFNL---ESRVEIEKSIQ 178
QY 206 THNRLKSLMKILSEV 220
DB 179 MEVLTAQKQREV 193

RESULT 5
ID CU045 HUMAN STANDARD; PRT; 233 AA.
AC G9NYE9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Protein C21orf45.
GN Name=C21orf45;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20237674; PubMed=10773462; DOI=10.1016/S0378-1119(00)00089-5;
RA Slavov D., Hattori M., Sakaki Y., Rosenthal A., Shimizu N.,
RA Minoshima S., Kudoh J., Yaspo M.-L., Ramser J., Reinhardt R.,

RA Taenzer S., Galgoczy P., Platzer M., Scharfe M., Nordsiek G.,
RA Bloecher H., Hellmann I., Khaitovich P., Paabo S., Reinhardt R.,
RA Zheng H.-J., Zhang X.-L., Zhu G.-F., Wang B.-F., Fu G., Ren S.-X.,
RA Zhao G.-P., Chen Z.-S., Lee Y.-S., Cheong J.-E., Choi S.-H., Wu K.-M.,
RA Liu T.-T., Haio K.-J., Tsai S.-P., Kim C.-G., Oota S., Kitano T.,
RA Kohara Y., Saitou N., Park H.-S., Wang S.-Y., Yaspo M.-L., Sakaki Y.,
RT "DNA sequence and comparative analysis of chimpanzee chromosome 22";
RL Nature 429:382-388 (2004).
CC -1- SIMILARITY: To S.pombe Spc970.12.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AL954205; CAH18576.1; -; Genomic DNA.
DR SEQUENCE 232 AA; 25832 MW; 6549BBB369245E7 CRC64;
SQ
Query Match 9.2%; Score 110.5; DB 1; Length 232;
Best Local Similarity 27.1%; Pred. No. 0.065;
Matches 52; Conservative 29; Mismatches 88; Indels 23; Gaps 8;
Qy 31 ASFTTSEWDVQVKGSSPLGAGPEAPGQPLPSWLOPERCAVFCQACQAVLADS 90
Db 49 ASMWSSMSSEASV-----ADMERARL--EAAAA-----EERPLVLCGCRRLPLGDS 94
Qy 91 VHLAWDLRS-LSGAVFVSRTNNVVLAPFLVIGSLKSTYLLFGSCGIPVGFHLY 149
Db 95 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKNGCVLTLCAGCSNLGVYR 152
Qy 150 STHAALALRGHFCSSDQVCYLL--KTKAIVNASMDIONVPLSEKIAELKEKIVLTH 207
Db 153 CTPKNLDYKDFLCVLSAEISYVLGSSEKQIV-SEDKELFNL---BSRVIEKSLQME 208
Qy 208 NRLKSLMKILSE 219
Db 209 DVLKALQKQWE 220
RESULT 8
Q5BLB1_BRARE PRELIMINARY; PRT; 207 AA.
AC Q5BLB1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE LOC553502 protein (Fragment).
GN Name=LOC553502;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Chatschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Beak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RG NIH MGC Project;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC090521; AAH90521.1; -; mRNA.
DR InterPro; IPR000074; ACoAL A4_E.
DR InterPro; IPR009090; D_Amino_Dept_C.
FT NON TER 1
SQ SEQUENCE 207 AA; 22941 MW; AF1A0ACC375A28C CRC64;
Query Match 9.1%; Score 109; DB 2; Length 207;
Best Local Similarity 25.5%; Pred. No. 0.077;
Matches 37; Conservative 28; Mismatches 66; Indels 14; Gaps 6;
Qy 76 AVFQACQAVLADSVHLDLS-RSLGAVFVSRTNNVVL-EAPFLVIGISLKGSTYN 133
Db 54 AVFMCCKCKLPIGDS--LSWAGSDDENNQIMLKITDNIIVGKEPFGSTKELGCLVNV 111
Qy 134 LLFCGSGCIPVGFHLYSTHAALALRGHFCSSDKVCYLLKTKAIVNASMDIONVPLS 193
Db 112 LT-CRGCSESLGVNVIPTPKLDCRSLFCFNVENIESYVVGSPG-QQPSLDREDKPT 169
Qy 194 -----EKIAELKEKIVLTHNRL 210
Db 170 LEYQDTVHQQWTEITSLAVIQRL 194
RESULT 9
Q54HI6_DICDI PRELIMINARY; PRT; 532 AA.
AC Q54HI6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80188417;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugand R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Davis P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loussegod H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum";
RL Nature 0-0-0 (2005).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF10100207; EAL62738.1; -; Genomic_DNA.
KW Hypothetical protein.

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 GN Putative gypsy-type retrotransposon protein.
 CN ORFNames=OSJNBa0096R22.5;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Sakaki C., Henry D., Oates R., Simmons J.,
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";
 RL Science 300:1566-1569 (2003).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.,
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC099400; AAL01599.1; -; Genomic_DNA.
 DR EMBL; AE017047; AAP01763.1; -; Genomic_DNA.
 DR Gramene; Q8S611; -;
 SQ SEQUENCE 361 AA; 38804 MW; F883B3B8B8FF45 CRC64;

Query Match 7.7%; Score 92; DB 2; Length 361;
 Best Local Similarity 25.3%; Pred. No. 6.4;
 Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;

Qy 2 AAQLRHSRCATPPRGFCGCTGTAIDQASFTTSMEDWTQVKGSSPLGAGLCABEPA 61
 Db 49 APTLPFRRAVRAAQQSGTSSAPVAST-----DVVVFGSREATPSG-PASDPV 102
 Qy 62 AGPOLP-----SW--LQPRCAVFOCAQCHAVLADSLRSLSGAVFVRVNNVL 115
 Db 103 AGRSPAAVLWEELQVEMGRLEAGA--RVIGREIAEARGLEHRM-----SELGN-- 152
 Qy 116 EAPFLVIGEGSLKSTYNLL--FCGSGI-----PVGFHYLTHAALAA-----RGH 161
 Db 153 ----LSEIRGSLR-VTYTGLHQLAGKGIKTIIPANDFEFLTSLAEALAAEIEFSKH 207
 Qy 162 FCLSSDKM-----VCYLLKTKAIVNASMDIQNV-----PLSEKIAELKEKI 203
 Db 208 AARIGEEMSNRIYIGACHILACVRLAH-PELDRLREILDQGEASDARKDVMEEVGLGKSV 266
 Qy 204 V 204
 Db 267 L 267

RESULT 13
 Q9CH87_LACLA
 ID Q9CH87_LACLA PRELIMINARY; PRT; 695 AA.
 AC Q9CH87;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Copper-potassium transporting ATPase B.
 GN Name=copB; OrderedLocusNames=LL0851;
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=ILL403;
 RC MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr-1697R;
 RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarne K.,
 RA Weissbach J., Ehrlich S.D., Sorokin A.;

"The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis ILL403.";
 RL Genome Res. 11:731-753 (2001).
 DR EMBL; AE006319; AAK04949.1; -; Genomic_DNA.
 DR PIR; C86731; C86731.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005524; F: ATP binding; IEA.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0004008; F: copper-exporting ATPase activity; IEA.
 DR GO; GO:0016820; F: hydrolase activity, acting on acid anhydrid. . .; IEA.
 DR GO; GO:0046872; F: metal ion binding; IEA.
 DR GO; GO:0046873; F: metal ion transporter activity; IEA.
 DR GO; GO:0008152; P: metabolism; IEA.
 DR GO; GO:0030001; P: metal ion transport; IEA.
 DR GO; GO:0015992; P: proton transport; IEA.
 DR InterPro; IPR006403; ATPase-IB1 Cu.
 DR InterPro; IPR006416; ATPase-IB hvy.
 DR InterPro; IPR001757; ATPase-IB E2.
 DR InterPro; IPR005834; Dehalo-like hydro.
 DR InterPro; IPR008250; E1-E2 ATPase_reg.
 DR Pfam; PF00122; E1-E2 ATPase; 1.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR TIGRFAMS; TIGR01511; ATPase-IB1 Cu; 1.
 DR TIGRFAMS; TIGR01525; ATPase-IB hvy; 1.
 DR TIGRFAMS; TIGR01494; ATPase-P-type; 2.
 DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 695 AA; 75981 MW; 1B4947C3C2A0FA0F CRC64;

Query Match 7.7%; Score 92; DB 2; Length 695;
 Best Local Similarity 25.3%; Pred. No. 15;
 Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;

Qy 101 LGAVFVRVNNVLAPFLVIGEGSLKSTYNLLFCGSGIIPVGFHYLTHAALAAALRG 160
 Db 101 LGTIIP-----FVSGTFPGGAKGKSRKPMMLITWGITVAY-AYSVYAIMSLNG 153
 Qy 161 HF-----CLSSDKMVCYLLKTKAIVNASB--MDI-----QNVPLSR-K 195
 Db 154 HMGNFWFELATLIVIMLIGHLEMKAIMGADKDLASLVPPKKAHLKSGKDVLSBLK 213
 Qy 196 IAELEKELVLTNRLKSLMKILSRVTPQS 225
 Db 214 VGDL--LLVKNKIPADGLILSEALVDES 241

RESULT 14
 Q69KE0_ORYSA
 ID Q69KE0_ORYSA PRELIMINARY; PRT; 628 AA.
 AC Q69KE0;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Myosin heavy chain-like.
 GN Name=P0023E10.17;
 OS Oryza sativa (japonica cultivar-group).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
 RT clone:P0023E10.17;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005934; BAD36600.1; -; Genomic_DNA.
 DR Gramene; Q69KE0; -;
 DR InterPro; IPR005829; Sug transporter.
 DR PROSITE; PS00216; SUGAR TRANSPORT 1; UNKNOWN_1.
 SQ SEQUENCE 628 AA; 68281 MW; 3230B108C6061027 CRC64;

Query Match 7.6%; Score 91; DB 2; Length 628;
Best Local Similarity 23.7%; Pred. No. 16;
Matches 66; Conservative 34; Mismatches 82; Indels 96; Gaps 17;
QY 8 HRSRCATPRG-----DPCGTERAIDQASFTTSMEDWTDVVKGSSPLGPAGLGA 57
DB 141 YRKAPTSPSRVARAKAQAQDGGTSSASPAVAST-----DVVVFGSREAAFGSL-A 194
QY 58 EEPAGPOLP-----SW--LQPERCAVFOC-----AQHAVLADSVH-----LAWDL 97
DB 195 SDLVAGPGSPAALVTWELQVEMGRLLGARGIGREIAEAAAAAASANERADRHLAHD 254
QY 98 S-----RSLGAV-----PSRVTVNVLEAPFLVIGIEGSLKSTYNLL--P 136
DB 255 ABAREDLKKRELVAQNRQOGLERHMRSELENN-----LSEIRDSLR-VITYTGLHQL 306
QY 137 CGSCGI-----PVGFHYLSTHAAAL-----RGHFCLSSDKM-----VCYLLKTK 177
DB 307 AGECKIKSTIPANPDEFSLTSSLAELAAAMEEIPSKHAARTRETSNGIYTGACHVLACL 366
QY 178 AIVNAEMDIQNV-----PLSEKIAELKEKIV 204
DB 367 RLVB-PELDRLDQGAASDARKDVMEEVGDGLESVL 403

RESULT 15

OB3693_TREPA PRELIMINARY; PRT; 597 AA.
AC OB3693;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Phosphoribosylglycinamide formyltransferase, putative.
GN OrderedLocusNames=TP0695;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Nichols;
RX MEDLINE=9832770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artiach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete";
RL Science 281:375-388(1998).
DR EMBL; AE001243; AAC65662.1; -; Genomic_DNA.
DR PIR; D71293; D71293.
DR TIGR; TP0695;
DR GO; GO:0016874; F:ligase activity; IEA.
DR InterPro; IPR00135; ATP-grasp.
DR InterPro; IPR011761; ATP_GRASP.
DR Pfam; PF02222; ATP-grasp; 1.
DR PROSITE; PS00975; ATP_GRASP; 1.
KW Complete proteome.
SQ SEQUENCE 597 AA; 63326 MW; 7B775117C9BFD6A6 CRC64;

Query Match 7.6%; Score 90.5; DB 2; Length 597;
Best Local Similarity 28.0%; Pred. No. 17;
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;
QY 12 CATPPRGDFCGGTERAID-----QASFT-----TSMEDTQVVKGSSPLGPA 53
DB 90 CALP--CHRLKATKATKTRMCACTFRALRCPRFTFLFDPSPFANDT-----PPGHA 140
QY 54 GLGAEEPAAGPQLPSMLOP--ERCAVFOC--AQHAVLADSVHAWDLRSILGAVVFSRVT 110

DB 141 RLCSHHSAGLSFPLVVKPTDNMGARGCTLAQCKDTLINACAVARQFSRS-----190
QY 111 NNVLLEAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFHYLSTHAAALRGHFCLSSDKW 170
DB 191 GRVIEE-EFIVGREFSLESG-----LIFDGT-----LYVT--ALA-----DRHI 225
QY 171 CY 172
DB 226 CF 227

Search completed: March 25, 2006, 08:03:11
Job time : 234 secs

Result No.	Query			ID	Description
	Score	Match	Length		
1	86	7.2	158	US-09-134-000C-6596	Sequence 6596, Appl
2	83	6.9	457	5268463-7	Patent No. 5268463
3	81.5	6.8	456	5432081-7	Patent No. 5432081
4	81.5	6.8	457	US-08-883-704A-6	Sequence 6, Appl
5	81.5	6.8	457	US-09-151-957-6	Sequence 6, Appl
6	80.5	6.7	1711	US-08-342-930-2	Sequence 2, Appl
7	78.5	6.6	325	US-09-018-576-3	Sequence 3, Appl
8	78.5	6.6	325	US-09-018-576-12	Sequence 12, Appl
9	78.5	6.6	325	US-09-248-137-3	Sequence 3, Appl
10	78.5	6.6	325	US-09-248-137-12	Sequence 12, Appl
11	78.5	6.6	452	US-09-771-161A-237	Sequence 237, Appl
12	78	6.5	190	US-09-252-991A-19049	Sequence 19049, A
13	77	6.4	607	US-09-537-682-1	Sequence 1, Appl
14	77	6.4	719	US-08-520-933-3	Sequence 3, Appl
15	77	6.4	719	US-09-285-040-3	Sequence 3, Appl
16	77	6.4	738	5262177-2	Patent No. 5262177
17	77	6.4	776	US-09-949-016-10829	Sequence 10829, A
18	76.5	6.4	282	US-09-845-713A-2	Sequence 2, Appl
19	76	6.3	351	US-09-902-540-13214	Sequence 13214, A
20	76	6.3	1001	US-10-104-047-2283	Sequence 2283, Ap
21	75.5	6.3	310	US-08-129-456A-36	Sequence 36, Appl
22	75.5	6.3	310	US-08-705-868-3	Sequence 3, Appl
23	75.5	6.3	310	US-09-123-615-3	Sequence 3, Appl
24	75.5	6.3	310	US-08-360-821B-35	Sequence 35, Appl
25	75.5	6.3	310	US-09-851-026-35	Sequence 35, Appl
26	75.5	6.3	1449	US-08-840-062-6	Sequence 6, Appl
27	75.5	6.3	1723	US-09-194-612A-31	Sequence 31, Appl

```

RESULT 2
5268463-7
; Patent No. 5268463
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: PLAAAT PROMOTER a-GLUCURONIDASE GENE
; :CONSTRUCT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,976
; FILING DATE: 08-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 119,102

```

; FILING DATE: 10-NOV-1987
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; SEQ ID NO: 7;
; LENGTH: 457
5268463-7

Query Match 6.9%; Score 83; DB 6; Length 457;
Best Local Similarity 25.2%; Pred. No. 2;
Matches 37; Conservative 30; Mismatches 54; Indels 26; Gaps 8;

QY 95 WDLSSLSGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 316 WSLPVALVALAIAASIGQGVTTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 373

QY 142 IPVGFHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASMDIQNV-----PLSE 194
Db 374 QAIGGS--SIPAFILGLSGYIANQVTPVINGIRTSIALVPCGFMLLAFVLIWFYPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KFKETVVEIDNRKKVQQQLISDIT 456

RESULT 3

5432081-7
; Patent No. 5432081
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: HOST CELLS TRANSFORMED WITH THE E. COLI
; GLUCORONIDE PERMEASE GENE
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,546
; FILING DATE: 15-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 447,976
; FILING DATE: 08-DEC-1989
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; APPLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; SEQ ID NO: 7;
; LENGTH: 456
5432081-7

Query Match 6.8%; Score 81.5; DB 6; Length 456;
Best Local Similarity 25.2%; Pred. No. 2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLSSLSGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 316 WSLPVALVALAIAASIGQGVTTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 373

QY 142 IPVGFHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASMDIQNV-----PLSE 194
Db 374 QAIG--GSIPAFILGLSGYIANQVTPVINGIRTSIALVPCGFMLLAFVLIWFYPLTD 430

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 431 K--KFKETVVEIDNRKKVQQQLISDIT 455

RESULT 4

US-08-882-704A-6
; Sequence 6, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,704A
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5879906tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 190106.404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-882-704A-6

Query Match 6.8%; Score 81.5; DB 1; Length 457;
Best Local Similarity 25.2%; Pred. No. 2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLSSLSGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIAASIGQGVTTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGFHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGLSGYIANQVTPVINGIRTSIALVPCGFMLLAFVLIWFYPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KFKETVVEIDNRKKVQQQLISDIT 456

RESULT 5

US-09-151-957-6
; Sequence 6, Application US/09151957
; Patent No. 6429292
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,957
; FILING DATE: 11-Sep-1998
; CLASSIFICATION: <Unknown>

Db 134 LISASGQLKIADFGFLARVFS---PDGSRILY-THQ--VATRSVGCIMGELLNGSLPLFPQGN 187
Qy 168 --KWCYLLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPPELTLPDYNKIS-FKEQVMP-----LEEVLDPV 236
Qy 225 S 225
Db 237 S 237

RESULT 8

US-09-018-576-12
; Sequence 12, Application US/09018576
; Patent No. 5968800
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,576
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-018-576-12

Query Match 6.6%; Score 78.5; DB 1; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.8;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
Qy 22 GGERAIDQASFTTSMEDTQVVGSSPLGPAGLG-----AEEPAAG 63
Db 43 GFNPQALREIKALQEME-DNQYVQLKAVFPHGGLVLAFFELMFLSDLAEVVRHQRPLAQ 101
Qy 64 POLPSWLOPERCAVFCQACQACHAVLADSVHLAWDLRSLSGAVVFSRVTVNNV-----LEAPF 119
Db 102 AQVKSILQMLKGV---AFCHA-----NNIVHRLKPNAL 133
Qy 120 LVGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCSSD----- 167
Db 134 LISASGQLKIADFGFLARVFS---PDGSRILY-THQ--VATRSVGCIMGELLNGSLPLFPQGN 187
Qy 168 --KWCYLLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPPELTLPDYNKIS-FKEQVMP-----LEEVLDPV 236
Qy 225 S 225

Db 237 S 237
RESULT 9
US-09-248-137-3
; Sequence 3, Application US/09248137
; Patent No. 6030788
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/248,137
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,576
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-248-137-3

Query Match 6.6%; Score 78.5; DB 2; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.8;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
Qy 22 GGERAIDQASFTTSMEDTQVVGSSPLGPAGLG-----AEEPAAG 63
Db 43 GFNPQALREIKALQEME-DNQYVQLKAVFPHGGLVLAFFELMFLSDLAEVVRHQRPLAQ 101
Qy 64 POLPSWLOPERCAVFCQACQACHAVLADSVHLAWDLRSLSGAVVFSRVTVNNV-----LEAPF 119
Db 102 AQVKSILQMLKGV---AFCHA-----NNIVHRLKPNAL 133
Qy 120 LVGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCSSD----- 167
Db 134 LISASGQLKIADFGFLARVFS---PDGSRILY-THQ--VATRSVGCIMGELLNGSLPLFPQGN 187
Qy 168 --KWCYLLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPPELTLPDYNKIS-FKEQVMP-----LEEVLDPV 236
Qy 225 S 225
Db 237 S 237

RESULT 10

RESULT 11
US-09-771-161A-237
; Sequence 237, Application US/09771161A
; Patent No. 6936450
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.

```

RESULT 12
US-09-252-991A-19049
; Sequence 19049, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19049
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19049

```

```

Query Match      6.5%; Score 78; DB 2; Length 190;
Best Local Similarity 32.9%; Pred. No. 1.9;
Matches 24; Conservative 6; Mismatches 29; Indels 14; Gaps 3;

4 QPLRHRSCRATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEPAAG 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 RPCRHTTRCVRPAAGGYPG---RAGDRAS-----VARSTGGLPAAPGRRPGHC- 139
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

QY 64 POLPSWQPERCA 76
DB 140 PAPPWRPRRCA 152

RESULT 13
US-09-537-682-1
; Sequence 1, Application US/09537682
; Patent No. 6303357
; GENERAL INFORMATION:
; APPLICANT: TAKUCHI, Kenichi
; APPLICANT: KOIDE, Yoshinao
; APPLICANT: NAKAMISHI, Yuji
; APPLICANT: SUZUKI, Satoru
; TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE, RECOMBINANT
; TITLE OF INVENTION: DNA, AND METHOD FOR PRODUCING MODIFIED
; TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE
; FILE REFERENCE: A20-121814C/KI
; CURRENT APPLICATION NUMBER: US/09/537,682
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Enterococcus faecium No. 6303357 7044
US-09-537-682-1

Query Match 6.4%; Score 77; DB 2; Length 607;
Best Local Similarity 23.2%; Pred. No. 15;
Matches 49; Conservative 30; Mismatches 76; Indels 56; Gaps 10;

QY 19 DFCGGERAIDQASFTSM-----EWDQVVGKSSPLGP-----AGLGAEEPAAGPQ 65
DB 51 DPAEET-----SSRSTKLHVGIGIRYKTFDVEV--ADTVGERAVVQGIAPHIPKPDPM 102
QY 66 LPSWQPERCAVQCAQCHAVLADSVHLAWDLRSLSGAVVFSRVN-----NVLEAPP 119
DB 103 LPIIYEDGATTFNMF-----SVKVAMDLYDKLANVTGKYENYLTTPPEVLEREPF 154
QY 120 LVGIGSLKGYSTNLLFCGSGGIPGVF-----HLYSTHAALAAHGHFCLSSDKMVCYLL 174
DB 155 LK--KEGLKA-----GVYLDFRNDRDLVIDNIKKAAEDGAYLVSRKAVGFLY 202
QY 175 KTKAIVNAEMDIQNVPLSEKIAELKEKIVLTHNRLKSL 205
DB 203 EGDQIVGVKARDL-----LTDEVIEIKSLVI 229

RESULT 14
US-08-520-933-3
; Sequence 3, Application US/08520933
; Patent No. 5981194
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L.
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/520,933
; FILING DATE: August 31, 1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Shona S. McDiarmid
; REGISTRATION NUMBER: 38,798
; REFERENCE/DOCKET NUMBER: 7685-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-364-7311
; TELEFAX: 416-361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-520-933-3

Query Match 6.4%; Score 77; DB 1; Length 719;
Best Local Similarity 23.7%; Pred. No. 19;
Matches 59; Conservative 25; Mismatches 87; Indels 78; Gaps 15;

QY 17 RGDGCGTERAIDQASFTSM-----DTQVVGKSSPLGPAGLGAEEPAAGP 64
DB 171 RGDSSG--EGVCDKSPLERYDYDGAFCRLAEGAGDVAFVKHSTVL-----ENTDGK 220
QY 65 QLPFWLQP-----ERCAVFOCAQCH--AVLADSVHLAWDLRSLSGAVVFSRV 109
DB 221 TLPFGWQALLSDPFELLCRDSRADVTWRQCHLARVPAHAVVVRADTD---GGLIF-RL 276
QY 110 TNNVLEAPFLVIGIEG-----LKGSTY---NLLFCGSCG--IPVGFHLYST---HAALAA 157
DB 277 LN-----EQRLFSEHSGSFQMFSEAYCGKOLLFKDSTSELVPIATQTYEAMLGHEYLHA 332
QY 158 LRGHFCLSSDKMVCYL-----LTKAIVNAEMDIQNVPLSEKIAELKEKIVLTHNRLKSL 213
DB 333 MKGLLC-DPNRLPPYLRWCVLSTPBQKGDH-----AVAFRRQRLKPE 375
QY 214 MKILSEVTP 222
DB 376 IQCVSAKSP 384

RESULT 15
US-09-285-040-3
; Sequence 3, Application US/09285040
; Patent No. 6455494
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L.
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 08:14:45 ; Search time 165 Seconds
(without alignments)
579.896 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAAQPLRHRSRCATPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	3	US-09-942-052-728
2	1198	100.0	229	3	US-09-942-052-729
3	1198	100.0	229	3	US-09-942-052-730
4	1198	100.0	229	3	US-09-942-052-731
5	138.5	11.6	164	3	US-09-942-052-707
6	138	11.5	229	3	US-09-942-052-704
7	113	9.4	233	3	US-09-890-688-82
8	113	9.4	233	4	US-10-408-765A-969
9	95.5	8.0	708	4	US-10-437-963-195427
10	92	7.7	361	4	US-10-437-963-154548
11	92	7.7	695	4	US-10-369-493-18389
12	89	7.4	664	3	US-09-780-525-2
13	88.5	7.4	435	4	US-10-437-963-186569
14	87.5	7.3	848	4	US-10-437-963-155606
15	86.5	7.2	669	4	US-10-282-122A-51802
16	85	7.1	337	4	US-10-156-761-10954
17	84.5	7.1	485	4	US-10-363-829-358
18	84	7.0	1905	4	US-10-424-599-223506
19	84	7.0	2902	4	US-10-437-963-139559
20	83.5	7.0	888	4	US-10-437-963-147897
21	83	6.9	663	4	US-10-282-122A-51673
22	82.5	6.9	124	4	US-10-425-115-216789
23	82.5	6.9	278	4	US-10-425-114-42310
24	82.5	6.9	325	4	US-10-663-896-2
25	82.5	6.9	493	5	US-10-213-974-36
26	82.5	6.9	502	4	US-10-424-599-226730
27	82.5	6.9	1043	4	US-10-369-493-9793

28 81.5 6.8 278 4 US-10-425-114-66220 Sequence 66220, A
29 81.5 6.8 457 4 US-10-195-518-6 Sequence 6, Appli
30 81.5 6.8 701 5 US-10-450-763-31748 Sequence 31748, A
31 81.5 6.8 1969 4 US-10-437-963-115342 Sequence 115342,
32 81.5 6.8 3067 5 US-10-631-467-1618 Sequence 1618, Ap
33 81 6.8 229 4 US-10-425-115-281194 Sequence 281194,
34 81 6.8 859 4 US-10-210-281-60 Sequence 60, Appl
35 80.5 6.7 282 5 US-10-450-763-51058 Sequence 51058, A
36 80.5 6.7 1711 5 US-10-029-345A-38 Sequence 38, Appl
37 80 6.7 305 4 US-10-183-116-59 Sequence 59, Appl
38 80 6.7 305 6 US-10-957-135-59 Sequence 59, Appl
39 80 6.7 305 6 US-11-083-611-59 Sequence 138084,
40 80 6.7 544 4 US-10-424-599-158084 Sequence 968, App
41 79.5 6.6 212 3 US-09-925-299-968 Sequence 968, App
42 79.5 6.6 212 3 US-09-925-299-968 Sequence 57058, A
43 79 6.6 196 5 US-10-450-763-44924 Sequence 44924, A
44 79 6.6 230 5 US-10-450-763-44924 Sequence 44919, A
45 79 6.6 271 5 US-10-450-763-44919

ALIGNMENTS

RESULT 1

US-09-942-052-728

; Sequence 728, Application US/09942052

; Publication NO. US20030170626A1

; GENERAL INFORMATION:

; APPLICANT: Raitano, Arthur B.

; APPLICANT: Paris, Mary

; APPLICANT: Hubert, Rene S.

; APPLICANT: Afar, Daniel

; APPLICANT: Ge, Wangmao

; APPLICANT: Challita-Eid, Pia M.

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3

; FILE REFERENCE: 51158-20028.00

; CURRENT APPLICATION NUMBER: US/09/942,052

; CURRENT FILING DATE: 2001-08-28

; PRIOR APPLICATION NUMBER: 60/228,432

; PRIOR FILING DATE: 2000-08-28

; NUMBER OF SEQ ID NOS: 744

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 728

; LENGTH: 229

; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIP5

; OTHER INFORMATION: Clone A protein

US-09-942-052-728

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPRGDFCGGTERAIDQASFTTSMENDTQVKGSSPLGPAIGAEPP 60

Db 1 MAAQPLRHRSRCATPRGDFCGGTERAIDQASFTTSMENDTQVKGSSPLGPAIGAEPP 60

Qy 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPPL 120

Db 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPPL 120

Qy 121 VGIEGSLKGYTYNLLFCGSCCIPVGFHYLSTHAALALRGHFCSSDRKWCYLLKTKAIV 180

Db 121 VGIEGSLKGYTYNLLFCGSCCIPVGFHYLSTHAALALRGHFCSSDRKWCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 2

US-09-942-052-729
; Sequence 729, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 729
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIP5
; OTHER INFORMATION: protein sequence
US-09-942-052-729

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAAGLGAEP 60
Db 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAAGLGAEP 60

Qy 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120
Db 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120

Qy 121 VGIEGSLKSTYNLLFCGSCGIPVGPHLYSTHAALALRGHFCILSSDKWVCYLLKTKAIV 180
Db 121 VGIEGSLKSTYNLLFCGSCGIPVGPHLYSTHAALALRGHFCILSSDKWVCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 3

US-09-942-052-730
; Sequence 730, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 730
; LENGTH: 229
; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: 85P1B3 protein

; OTHER INFORMATION: sequence

US-09-942-052-730

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAAGLGAEP 60
Db 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAAGLGAEP 60

Qy 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120
Db 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120

Qy 121 VGIEGSLKSTYNLLFCGSCGIPVGPHLYSTHAALALRGHFCILSSDKWVCYLLKTKAIV 180
Db 121 VGIEGSLKSTYNLLFCGSCGIPVGPHLYSTHAALALRGHFCILSSDKWVCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 4

US-09-942-052-731
; Sequence 731, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 731
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-942-052-731

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAAGLGAEP 60
Db 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAAGLGAEP 60

Qy 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120
Db 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120

Qy 121 VGIEGSLKSTYNLLFCGSCGIPVGPHLYSTHAALALRGHFCILSSDKWVCYLLKTKAIV 180
Db 121 VGIEGSLKSTYNLLFCGSCGIPVGPHLYSTHAALALRGHFCILSSDKWVCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 5
US-09-942-052-707
; Sequence 707, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Splice variant 1
; OTHER INFORMATION: open reading frame 3 peptide sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (44)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (50)..(51)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (67)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (93)
; OTHER INFORMATION: Unknown amino acid or stop
US-09-942-052-707

Query Match 11.6%; Score 138.5; DB 3; Length 164;
Best Local Similarity 82.1%; Pred. No. 1.2e-05;
Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 162 FCLSSDKWVCYLKTKAIVNASEMDIQNVPLSEKIAELK 200
Db 2 FFLSS-----YLLKTKAIVNASEMDIQNVPLSEKIAEVK 35

RESULT 6
US-09-942-052-704
; Sequence 704, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28

; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 704
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3 peptide
; OTHER INFORMATION: sequence
US-09-942-052-704

Query Match 11.5%; Score 138; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 YLLKTKAIVNASEMDIQNVPLSEKIAELK 200
Db 1 YLLKTKAIVNASEMDIQNVPLSEKIAELK 29

RESULT 7
US-09-890-688-82
; Sequence 82, Application US/09890688
; Publication No. US20030144475A1
; GENERAL INFORMATION:
; APPLICANT: Seishi KATO
; APPLICANT: Chikashi EGUCHI
; APPLICANT: Mihoro SAEKI
; TITLE OF INVENTION: Human Proteins and cDNAs thereof
; FILE REFERENCE: 2001-11024/WMC/00653
; CURRENT APPLICATION NUMBER: US/09/890,688
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: JP 11-346863
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 11-34684
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 2000-31062
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: JP 2000-34091
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-34090
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-35829
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-35899
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-71161
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: JP 2000-160851
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-890-688-82

Query Match 9.4%; Score 113; DB 3; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0075;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

Qy 31 ASFTTSMEWDVTQVKGSSPLGPAAGLGAEPAGPQLPSWLQPERCAVFCACQACHAVLADS 90
Db 49 ASMWSSMSDASV-----ADMERAAQL--EEFAAAAE-----ERPLVFLSCGRRPLGDS 95

Qy 91 VHLAWDLSS- SLGAVVFSRVTVNNVLEAPFLVGIEGSLKSTVNLFCGSCGIPVGPHLY 149
Db --LSWVASOEDTNCILLRCVSCNVSDVKEQKLSKREKNGCVLETCCAGCCSLNIGVYR 153

Qy 150 STHAALAAALRGHFCILSSDRWCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
Db 154 CTPKNLDYKRDLCFLSLVEAIESVYVLSGSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209

QY 208 NLRKSLMKILSE 219
: ||: ||: ||
Db 210 DVLKALQMKLWE 221

RESULT 8

US-10-408-765A-969

; Sequence 969, Application US/10408765A

; Publication No. US20040101874A1

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408, 765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 969

; LENGTH: 233

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-969

Query Match 9.4%; Score 113; DB 4; Length 233;

Best Local Similarity 27.1%; Pred. No. 0.0075;

Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTSMEDTQVVKGSSPLGPAAGAEPAAGPOLPSWLPQPCAVFQCAOCHAVLADS 90

Db 49 ASNWSMSDASV-----ADMERQL--EEEAARAE-----EPLVFLSGCGRPLGDS 95

QY 91 VHLAWLDR--SLGAVFVRVNNVLEAPPLVIGESLKGSTYNNLLFCGSCGIPVGHLY 149

Db 96 --LSWVASQEDTNCILLRCVSCNVSDVDEKQKSKREKENGCVLETLCCAGCSLNLGVYR 153

QY 150 STHAALAAALGHFLCSDDKMKVCYL--KTKAIVNASEMDIONVPLSEKIAELKEKIVLTH 207

Db 154 CTPLNDYKRDPLCLSVESVIGSSSEKQIV--SEDKELFNL---ESRVEIERKSLTQME 209

QY 208 NLRKSLMKILSE 219

: ||: ||: ||

Db 210 DVLKALQMKLWE 221

RESULT 9

US-10-437-963-195427

; Sequence 195427, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437, 963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 195427

; LENGTH: 708

; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91377C.1.pep
US-10-437-963-195427

Query Match 8.0%; Score 95.5; DB 4; Length 708;

Best Local Similarity 24.5%; Pred. No. 2.1;

Matches 71; Conservative 32; Mismatches 86; Indels 101; Gaps 17;

QY 6 LRHRSRCATPPRGDF-----CGGTERAIDQASFTTSMEW 39

Db 225 LRHRRKANTESADSPPPRRQSIIVTIGKEARAKAARAKSGTSA-----SPTTVST 279

QY 40 DTQVVKGSSPLGPAAGAEPAAGPOLP-----SW--LQPERCAVF-----OCAQCH 84

Db 280 DVVVPVVGSOBVTGSG--PISDPAGGSLPEAVLTWEEQLQVEMGRLLKAGAGIGREISEAR 338

QY 85 AVLADS VHLAWLDRSLG-----AVFVRVNNVLE-----APLVCIEGLSK 128

Db 339 AETAAANALAERLVRELAAREDLTKMRELVAGNEROSKGLDRMSLGNLSIRGSLR 398

QY 129 GSTYNLL--FCGSCG---IPVG---PHLYSTHAALAAALRG-----HFCLSSDKM----- 169

Db 399 -VYTGHLQLAKECGVKSTIPVNLDRFSLTSSLAELATANGEIPSKHTSRIAEETNGIY 457

QY 170 --VCYLLKTKAIVNAS--EMDIONV-----PLSEKIAELKEKIV 204

Db 458 TGVCHVL---ACVRLSRPELDLREILDQGAASDTRKEVMEVSDLSGSLV 504

RESULT 10

US-10-437-963-154548

; Sequence 154548, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437, 963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 154548

; LENGTH: 361

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_54398C.1.pep

US-10-437-963-154548

Query Match

Best Local Similarity 25.3%; Pred. No. 1.9;

Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;

QY 2 AAQPLHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSSPLGPAAGAEPA 61

Db 49 APTPLFRRAVRAKAAQDGGSGTSSAPAVAST-----DVVVVPGSGREATPFG--PASDPV 102

QY 62 AGPOLP-----SW--LQPERCAVFQCAOCHAVLADS VHLAWLDRSLGAVFVRVNNVVL 115

Db 103 AGRGSPAIVLSWEELQVEMGRLLLEACA--RVIGRETAERGLEHRM-----SELGNN--- 152

QY 116 EAPFLVIGIEGLSKGSTYNLL--FCGSCGI-----PVGPHLYSTHAALAAAL-----RGH 161

Db 153 ----LSEIRGSLR-VYTTGLHQLAGKCGIKSTIPANPDEFSLTSSLAELAAAMEEIPSKH 207

QY 162 FCLSDKM-----VCYLLTKTKAIVNASEMDIQNV-----PLSEKIAELKEKI 203
Db 208 AARIGEEMSNRIYTGACHILACVRLAH-PELDREILDQGEASDARKDVMBEVGLGKSV 266
QY 204 V 204
Db 267 L 267

RESULT 11

US-10-369-493-18389
; Sequence 18389, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18389
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Lactococcus lactis
US-10-369-493-18389

Query Match 7.7%; Score 92; DB 4; Length 695;
Best Local Similarity 25.3%; Pred. No. 4.7;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
QY 101 LGAVFSRVNNVLEAPFLVIGESLKGSTYNLLFCGSCGIPVGFHLYSTHAALRG 160
Db 101 LGTIIP-----FYSGGTFFSGAGELKSRKPPAMMLITWGITVAY-AYSVIATIMSLG 153
QY 161 HF-----CLSSDKMVCYLLTKTKAIVNASE-MDI-----QNVPLSE-K 195
Db 154 HMGNFWEFELATLIVMLIGHLEMKAIMGAGDALDOLASLVKPKAHLKSGKDVSELSK 213
QY 196 IAEKKEKIVLTHRLKLSIMKILSEVTPDQS 225
Db 214 VGDL--LLVKNENKIPADGLILSEALVDES 241

RESULT 12

US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US20020004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Hurle
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: PHARI, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT APPLICATION NUMBER: US/09/780,525
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS

US-09-780-525-2

Query Match 7.4%; Score 89; DB 3; Length 664;
Best Local Similarity 23.0%; Pred. No. 8.9;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
QY 8 HRSRCATPPRGDFC-----CGTERAIDQA-----SFTTSMEMDQVVKGSPL----- 50
Db 434 YRRQAQPPH---CPAPEGEPGAPALGDAPSTSVSLTTAVQDYVCPQLQGSALCTCCFPQ 490
QY 51 ---GPAGLGAEEPAAGPOLPSWLPQRCVAFQCAQCHAVLADSVHLANDLSR----- 99
Db 491 PMPDRRAERQDPRVAQ-----OCAVC-----LQPFCHLYWGCTRTGCGCLA 534
QY 100 -----SLGAVVFSRVNNVLEAPFLVIGESLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
Db 535 PFCELNLGDKLGDVLANNSYESDILKNYLAT-RGLTWKNM-----TES 578
QY 154 ALAALRGHFCLS-----SDKMVCYLLTKTKAIVNASEMDIQNVPLSE 194
Db 579 LVALQRGVFLLSRVYRTGDTVLVCYCCGLRSFRELTYQYRONIPASE 624

RESULT 13

US-10-437-963-186569
; Sequence 186569, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 186569
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83357C.1.pep
US-10-437-963-186569

Query Match 7.4%; Score 88.5; DB 4; Length 435;
Best Local Similarity 22.9%; Pred. No. 5.6;
Matches 50; Conservative 27; Mismatches 82; Indels 59; Gaps 12;
QY 2 AAQPL--RH--RSRCATPPRGDFCGTERAIDQAFTTSMEMDQVVKGSPLGPAGLGA 57
Db 41 APEPLSCRHGRHLRCAVD-----GGAGRETERPSPAPQ-----RESPSGSLGAAL 88
QY 58 EEPAGPOLPSWLPQRCVAFQCAQCHA--VLADSVHLAWDLRSLSGAVVFSRVNNVVL 115
Db 89 EDPSPQPGVPLILLP-----LCCRCYAKEICSEYVVRITDLVNI-----LNSNAIS 135
QY 116 EAPFLVIGESLKGSTYNLLFCGSCGIPVGFHLYSTHAALARGHFCLSDKRVVCLLK 175
Db 136 EGPFF-----SMRKARFLGSAS---AFSVKQTEWP-----CATTSKKCY-LQ 173
QY 176 TKAIVNASEMDIQN-----VPLSEKIAELKEKIVLTH 207
Db 174 NGSFGGITEDEQSSLYNFLYPSKELLPDDKEMSIIDH 211

RESULT 14

US-10-437-963-155606

Sequence 155606, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 155606
LENGTH: 848
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(848)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_55353C.1.pep
US-10-437-963-155606

Query Match 7.3%; Score 87.5; DB 4; Length 848;
Best Local Similarity 22.1%; Pred. No. 18;
Matches 65; Conservative 23; Mismatches 109; Indels 97; Gaps 13;

QY	DB	16	PRGDFCGGTERAIDQASFTTSMEDWTQVVGKSSPLGPAGLGAEPP	----- 60
QY	DB	85	PRGFRLDGTQGVQRTASCHLTROLTDPLFKSGGCLGPRGRKRNTPPPPLGNQBEGRHL	144
QY	QY	61	AAGPQLPSWLOPERCAVFOCAQCH	----- AVLADSVHLAWDLSR ----- 99
DB	DB	145	TLGPDVPEGARPRGRCLCLPQASGHDTPGIPSTSLVNLKRTRIEDVVATMDISEANEG	204
QY	QY	100	--SLGAVV-FSRVTNNV-----VLEAPLVGIEGSLK-GSTYNLLFCG	----- 138
DB	DB	205	YVSCGSVIENSROMKAARAGVRAQSXPACXGIDPTVGRGVSV--LLFTARMGEGLCHEMTT	262
QY	QY	139	-----SCGIPV-----GFHLYSTH-----AALAAALRGHFCISSDRKVCYL	173
DB	DB	263	GSKAMEGCVKMCVWGMPWLAWDELGRSGYOGFLNHRRLPVKAVLDGHLCEEDAVNCLP	322
QY	QY	174	LKTKAIVNASEMDIONVPLSEKIAELKEKIVLTNRNLKSLMKILSEVTPDOSKP	227
DB	DB	323	KLRKTVGSASE-----AKEAVKPAVKQEKKIRS-VKVLVLVS-DLSLP	363

RESULT 15
US-10-282-122A-51802
Sequence 51802, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 08:15:45 ; Search time 24 Seconds
(without alignments)
281.425 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAAPLHRSCATPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174695 seqs, 29494374 residues

Total number of hits satisfying chosen parameters: 174695

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pap.*
- 2: /SID55/ptodata/1/pubpaa/US06 NEW PUB.pap.*
- 3: /SID55/ptodata/1/pubpaa/US07 NEW PUB.pap.*
- 4: /SID55/ptodata/1/pubpaa/PCT NEW PUB.pap.*
- 5: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pap.*
- 6: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pap.*
- 7: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pap.*
- 8: /SID55/ptodata/1/pubpaa/US60 NEW PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	7.2	584	7	US-11-096-568A-12723
2	86	7.2	619	7	US-11-096-568A-12722
3	86	7.2	644	7	US-11-096-568A-12721
4	80.5	6.7	1711	7	US-11-143-984A-38
5	79.5	6.6	227	7	US-11-096-568A-19014
6	79.5	6.6	278	7	US-11-096-568A-19013
7	79.5	6.6	282	7	US-11-096-568A-19012
8	79.5	6.6	306	7	US-11-172-740-327
9	79	6.6	482	7	US-11-229-371-87
10	79	6.6	482	7	US-11-229-371-177
11	79	6.6	482	7	US-11-228-923-87
12	79	6.6	482	7	US-11-228-923-177
13	79	6.6	482	7	US-11-228-875-87
14	79	6.6	482	7	US-11-228-875-177
15	79	6.6	1705	7	US-11-143-984A-37
16	77.5	6.5	462	7	US-11-087-099-12028
17	76.5	6.4	400	7	US-11-229-371-2
18	76.5	6.4	400	7	US-11-228-923-2
19	76.5	6.4	400	7	US-11-228-875-2
20	76	6.3	1001	7	US-11-072-512-2283
21	75.5	6.3	8635	7	US-11-205-109-15
22	75	6.3	295	7	US-11-172-740-323
23	75	6.3	458	7	US-11-069-642-3
24	74.5	6.2	267	7	US-11-087-099-9317
25	74.5	6.2	467	7	US-11-096-568A-11427

Sequence 4, Appli
Sequence 9, Appli
Sequence 219, App
Sequence 2, Appli
Sequence 73, Appli
Sequence 7267, Ap
Sequence 26, Appli
Sequence 48, Appli
Sequence 30876, A
Sequence 30875, A
Sequence 30874, A
Sequence 12205, A
Sequence 24485, A
Sequence 24484, A
Sequence 24483, A
Sequence 3758, Ap
Sequence 752, App
Sequence 90, Appl
Sequence 3468, App
Sequence 758, App

26 74.5 6.2 3689 7 US-11-075-185-4
27 74 6.2 323 7 US-11-140-625-9
28 74 6.2 428 6 US-10-644-807-219
29 74 6.2 428 7 US-11-127-622-2
30 74 6.2 428 7 US-11-127-877-73
31 74 6.2 745 7 US-11-087-099-7267
32 74 6.2 3063 7 US-11-186-284-26
33 74 6.2 7102 7 US-11-143-980-48
34 73 6.1 256 7 US-11-096-568A-30875
35 73 6.1 274 7 US-11-096-568A-30876
36 73 6.1 327 7 US-11-096-568A-30874
37 72.5 6.1 378 7 US-11-087-099-12205
38 72.5 6.1 532 7 US-11-096-568A-24485
39 72.5 6.1 638 7 US-11-096-568A-24484
40 72.5 6.1 644 7 US-11-096-568A-24483
41 72 6.0 205 7 US-11-072-512-3758
42 72 6.0 388 7 US-11-055-822-752
43 71.5 6.0 1006 7 US-11-203-251A-90
44 71 5.9 229 7 US-11-072-512-3468
45 71 5.9 273 6 US-10-330-773-758

ALIGNMENTS

RESULT 1
US-11-096-568A-12723
; Sequence 12723, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12723
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(584)
; OTHER INFORMATION: Ceres Seq. ID no. 14303176
US-11-096-568A-12723

Query Match 7.2%; Score 86; DB 7; Length 584;
Best Local Similarity 21.5%; Pred. No. 1.3;
Matches 47; Conservative 23; Mismatches 69; Indels 80; Gaps 9;
QY 72 PERCAVFOCAQCHAV-----LADSVHLA-WDLRSRL-----GAVVFS 107
DB 87 PIPAEIRNCTELRAIVLRANYLOGGIPPGIGDLTHLTILDLSNLRGAIPATIGSLTHL 146
QY 108 RVTNNVVLEAPFL-----VGIEGSLKSTY--NLPCG-----SCGIPGVFHYLSTHA- 153
DB 147 RFLN---LSTNFFSGEIPNVGLTFKSSSYVGNLELCGLPIQKGCRTGLGPAVLPHSD 203
QY 154 -----ALALRGHFCLSSDKMVCYLLKTKAIV 180
DB 204 PLSSSGVSPITNNKTSFHLNGVVGISMSTMAIALVAVLGL-----WVCLLSRKNGV 257
QY 181 NASEMDIQNVLPSEKIAELKEKIVLTHNRLKSLMKILSE 219
DB 258 NYEKMDQIVPDGAKLVTVQWNLPSYSSGEIRLELLDE 296

RESULT 2
US-11-096-568A-12722
; Sequence 12722, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 12721
LENGTH: 619
TYPE: PR1
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(619)
OTHER INFORMATION: Ceres Seq. ID no. 14303175
US-11-096-568A-12721

Query Match 7.2%; Score 86; DB 7; Length 619;
Best Local Similarity 21.5%; Pred. No. 1.4; Mismatches 69; Indels 80; Gaps 9;
Matches 47; Conservative 23;

QY 72 PERCAVFOCAQCHAV-----LADSVHLA-WDLRSRL-----GAVVFS 107
DB 122 PIPAEIRNCTELRAIVLRANVYQGGIPPGIGDLTLTILDLSSNLLRGAIPATIGSLTHL 181
QY 108 RVTNNVLEAPFL-----VGIEGSLKSTY--NLLFCG-----SCGIPVGFHLYSTHA- 153
DB 182 RPLN---LSTNFFSGEIPNVGVLTGPKSSSYVGNLELCGLPIQKGCRTGLGFPVAVLPHSD 238
QY 154 -----ALAAIRGHFCLSSDKMVCYLLKTKAIV 180
DB 239 PLSSSGVSPITNNKNTSHFLNGVVGISMSTMAIALVAVLGFL-----WVCLLSRKNGV 292

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSE 219
DB 293 NYEKMDKQTVPDGAKLVTYQWNLPSYSGEIRRLLELDE 331

RESULT 3
US-11-096-568A-12721
Sequence 12721, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 12721
LENGTH: 644
TYPE: PR1
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(644)
OTHER INFORMATION: Ceres Seq. ID no. 14303174
US-11-096-568A-12721

Query Match 7.2%; Score 86; DB 7; Length 644;
Best Local Similarity 21.5%; Pred. No. 1.5; Mismatches 69; Indels 80; Gaps 9;
Matches 47; Conservative 23;

QY 72 PERCAVFOCAQCHAV-----LADSVHLA-WDLRSRL-----GAVVFS 107
DB 147 PIPAEIRNCTELRAIVLRANVYQGGIPPGIGDLTLTILDLSSNLLRGAIPATIGSLTHL 206
QY 108 RVTNNVLEAPFL-----VGIEGSLKSTY--NLLFCG-----SCGIPVGFHLYSTHA- 153
DB 207 RPLN---LSTNFFSGEIPNVGVLTGPKSSSYVGNLELCGLPIQKGCRTGLGFPVAVLPHSD 263
QY 154 -----ALAAIRGHFCLSSDKMVCYLLKTKAIV 180

DB 264 PLSSSGVSPITNNKNTSHFLNGVVGISMSTMAIALVAVLGFL-----WVCLLSRKNGV 317
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSE 219
DB 318 NYEKMDKQTVPDGAKLVTYQWNLPSYSGEIRRLLELDE 356

RESULT 4
US-11-143-984A-38
Sequence 38, Application US/11143984A
Publication No. US20060014180A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL HUMAN PHOSPHATASES
FILE REFERENCE: D0072 DIV1
CURRENT APPLICATION NUMBER: US/11/143,984A
CURRENT FILING DATE: 2005-06-02
PRIOR APPLICATION NUMBER: US 60/256,868
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: US 60/280,186
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/287,735
PRIOR FILING DATE: 2001-05-01
PRIOR APPLICATION NUMBER: US 60/295,848
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/300,465
PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.2
SEQ ID NO 38
LENGTH: 1711
TYPE: PR1
ORGANISM: Rattus norvegicus
US-11-143-984A-38

Query Match 6.7%; Score 80.5; DB 7; Length 1711;
Best Local Similarity 29.6%; Pred. No. 20; Mismatches 21; Conservative 8; Mismatches 17; Indels 25; Gaps 1;
Matches 21;

QY 36 SMEWDQVYKGSPLGPAAGLGAEPAPQPLPSWLOPERCAVFOCAQCHAVLADSVHLAW 95
DB 753 SMOAGSAVNLAWSGFLGQGA-----CHAQLSDAGHLWS 787

QY 96 DLRSRLGAVVF 106
DB 788 EQPLKLGQELF 798

RESULT 5
US-11-096-568A-19014
Sequence 19014, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 19014
LENGTH: 227
TYPE: PR1
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(227)
OTHER INFORMATION: Ceres Seq. ID no. 12368187
FEATURE:
NAME/KEY: misc feature
LOCATION: (122)..(122)
OTHER INFORMATION: Xaa is any aa, unknown or other

FEATURE:
NAME/KEY: misc feature
LOCATION: (126)..(126)
OTHER INFORMATION: Xaa is any aa, unknown or other
FEATURE:
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-19014

Query Match 6.6%; Score 79.5; DB 7; Length 227;
Best Local Similarity 24.5%; Pred. No. 1.6;
Matches 24; Conservative 11; Mismatches 40; Indels 23; Gaps 4;

Qy 10 SRCATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAGLGAEEPAAGPQLPSW 69
Db 118 SGCCXPP--TYCNFT-----YTXGTEW-TRTAAGSAP-----AGRDCSAW 154

Qy 70 LQPERCAVFOCAOCHAVLADSVHLAWDLRSLSGAVVFS 107
Db 155 GNEDDDLCCGQCKAGVVDALKRDWKRAAIVNVVILS 192

RESULT 6

US-11-096-568A-19013
Sequence 19013, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:

APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 19013
LENGTH: 278
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:

NAME/KEY: misc feature
LOCATION: (1)..(278)
OTHER INFORMATION: Xaa is any aa, unknown or other
FEATURE:
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any aa, unknown or other
FEATURE:
NAME/KEY: misc feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-19013

Query Match 6.6%; Score 79.5; DB 7; Length 278;
Best Local Similarity 24.5%; Pred. No. 2.1;
Matches 24; Conservative 11; Mismatches 40; Indels 23; Gaps 4;

Qy 10 SRCATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAGLGAEEPAAGPQLPSW 69
Db 169 SGCCXPP--TYCNFT-----YTXGTEW-TRTAAGSAP-----AGRDCSAW 205

Qy 70 LQPERCAVFOCAOCHAVLADSVHLAWDLRSLSGAVVFS 107
Db 206 GNEDDDLCCGQCKAGVVDALKRDWKRAAIVNVVILS 243

RESULT 7

US-11-096-568A-19012
Sequence 19012, Application US/11096568A

Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 19012
LENGTH: 282
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(282)
OTHER INFORMATION: Ceres Seq. ID no. 12368185
FEATURE:
NAME/KEY: misc feature
LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any aa, unknown or other
FEATURE:
NAME/KEY: misc feature
LOCATION: (181)..(181)
OTHER INFORMATION: Xaa is any aa, unknown or other
FEATURE:
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-19012

Query Match 6.6%; Score 79.5; DB 7; Length 282;
Best Local Similarity 24.5%; Pred. No. 2.2;
Matches 24; Conservative 11; Mismatches 40; Indels 23; Gaps 4;

Qy 10 SRCATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAGLGAEEPAAGPQLPSW 69
Db 173 SGCCXPP--TYCNFT-----YTXGTEW-TRTAAGSAP-----AGRDCSAW 209

Qy 70 LQPERCAVFOCAOCHAVLADSVHLAWDLRSLSGAVVFS 107
Db 210 GNEDDDLCCGQCKAGVVDALKRDWKRAAIVNVVILS 247

RESULT 8

US-11-172-740-327
Sequence 327, Application US/11172740
Publication No. US20060057724A1
GENERAL INFORMATION:

APPLICANT: MASCIA, Peter
APPLICANT: ALEXANDROV, Nickolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
FILE REFERENCE: 2750-1602PUS2
CURRENT APPLICATION NUMBER: US/11/172,740
CURRENT FILING DATE: 2005-06-30
PRIOR APPLICATION NUMBER: 60/583,621
PRIOR FILING DATE: 2004-06-30
PRIOR APPLICATION NUMBER: 60/584,829
PRIOR FILING DATE: 2004-06-30
PRIOR APPLICATION NUMBER: 60/584,800
PRIOR FILING DATE: 2004-06-30
NUMBER OF SEQ ID NOS: 2523
SEQ ID NO 327
LENGTH: 306
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(306)
OTHER INFORMATION: Ceres CLONE ID no. 1459706
FEATURE:


```
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for delaying flowering time
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for increasing chlorophyll and photosynthetic cap
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making biomass and foli
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making smaller plants
; US-11-172-740-327

Query Match      6.6%; Score 79.5; DB 7; Length 306;
Best Local Similarity 23.7%; Pred. No. 2.4;
Matches 46; Conservative 30; Mismatches 95; Indels 23; Gaps 10;

Qy 42 QVKGSSPLGAGAGAEPAAGPOLPSWLOPERCAVFOCAQ-CHAVLADSVHLAWDLRS 100
Db 67 QAVSKSHSLGHAASAPADHAASGVKTRMCTKYNATGCK--FGDKCHFAHG-ERE 123
Qy 101 LGAVVFSRVNNVLEAPFLVGIEGSLKSTYNLLFCGSCGIPVG-FHLYST-----HAA 154
Db 124 LGRPPSSYMQESSYAPPM-----GGRYGGRHEPPLPASMGPAGNFGASSTCKVSYDAA 178
Qy 155 LA-ALRGHFLSSDKVCYLLKTKAIV--NASEMDIQNVPLSEKIEKIVLTHNRLK 211
Db 179 LAGGIKGGVNT-KQICRTGVKLSIRDHESNFDLKNIELEGSFDQIKQ-----ANDMVR 233
Qy 212 SLMKILSEVTPDOS 225
Db 234 DLIASISASTPSKN 247

RESULT 9
US-11-229-371-87
; Sequence 87, Application US/11229371
; Publication No. US20060021092A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghasseman, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007CIP2CON1
; CURRENT APPLICATION NUMBER: US/11/229,371
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; TYPE: PRT
; LENGTH: 482
; ORGANISM: Unknown Organism

; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for delaying flowering time
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for increasing chlorophyll and photosynthetic cap
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making biomass and foli
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making smaller plants
; US-11-172-740-327

Query Match      6.6%; Score 79.5; DB 7; Length 306;
Best Local Similarity 23.7%; Pred. No. 2.4;
Matches 46; Conservative 30; Mismatches 95; Indels 23; Gaps 10;

Qy 42 QVKGSSPLGAGAGAEPAAGPOLPSWLOPERCAVFOCAQ-CHAVLADSVHLAWDLRS 100
Db 67 QAVSKSHSLGHAASAPADHAASGVKTRMCTKYNATGCK--FGDKCHFAHG-ERE 123
Qy 101 LGAVVFSRVNNVLEAPFLVGIEGSLKSTYNLLFCGSCGIPVG-FHLYST-----HAA 154
Db 124 LGRPPSSYMQESSYAPPM-----GGRYGGRHEPPLPASMGPAGNFGASSTCKVSYDAA 178
Qy 155 LA-ALRGHFLSSDKVCYLLKTKAIV--NASEMDIQNVPLSEKIEKIVLTHNRLK 211
Db 179 LAGGIKGGVNT-KQICRTGVKLSIRDHESNFDLKNIELEGSFDQIKQ-----ANDMVR 233
Qy 212 SLMKILSEVTPDOS 225
Db 234 DLIASISASTPSKN 247

RESULT 9
US-11-229-371-87
; Sequence 87, Application US/11229371
; Publication No. US20060021092A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghasseman, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007CIP2CON1
; CURRENT APPLICATION NUMBER: US/11/229,371
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; TYPE: PRT
; LENGTH: 482
; ORGANISM: Unknown Organism
```

```
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Description of Unknown Organism: Wigwag gene
; US-11-229-371-87

Query Match      6.6%; Score 79; DB 7; Length 482;
Best Local Similarity 26.3%; Pred. No. 5.1;
Matches 47; Conservative 25; Mismatches 49; Indels 58; Gaps 13;

Qy 28 IDQASFTTSMEDWTQ-----VVKSSPLGP--AGLGAEEPAAGPOLPSWLOPERCAVFOCAQ 82
Db 72 VSTQKYMIEIQRDKQDLMLKGLRQLGPFQFSSLDANRP-----WL-----111
Qy 83 CHAVLADSVHLAWDLRSIGAVVFSRVNNVLEAPFLVGIEGSLKSTYNLLFCGSCG- 141
Db 112 CYWIL-HSIAL-----LGETVDDLESNAI---DFLGRCCQSEGG-----YGGGPGQ 154
Qy 142 IPVGPHLYSTHA---ALAAALRGHFLCSS---DKMVCYLLKTK-----AIVNASEMDIQ 188
Db 155 LP---HLATTYAAVNALVTLLGGDKALSSINREKMSCFLRMKDTSGGFRHMDMGEMDVR 210

RESULT 10
US-11-229-371-177
; Sequence 177, Application US/11229371
; Publication No. US20060021092A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghasseman, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007CIP2CON1
; CURRENT APPLICATION NUMBER: US/11/229,371
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 177
; TYPE: PRT
; LENGTH: 482
; ORGANISM: Arabidopsis sp.

US-11-229-371-177

Query Match      6.6%; Score 79; DB 7; Length 482;
Best Local Similarity 26.3%; Pred. No. 5.1;
Matches 47; Conservative 25; Mismatches 49; Indels 58; Gaps 13;

Qy 28 IDQASFTTSMEDWTQ-----VVKSSPLGP--AGLGAEEPAAGPOLPSWLOPERCAVFOCAQ 82
Db 72 VSTQKYMIEIQRDKQDLMLKGLRQLGPFQFSSLDANRP-----WL-----111
Qy 83 CHAVLADSVHLAWDLRSIGAVVFSRVNNVLEAPFLVGIEGSLKSTYNLLFCGSCG- 141
Db 112 CYWIL-HSIAL-----LGETVDDLESNAI---DFLGRCCQSEGG-----YGGGPGQ 154
Qy 142 IPVGPHLYSTHA---ALAAALRGHFLCSS---DKMVCYLLKTK-----AIVNASEMDIQ 188
Db 155 LP---HLATTYAAVNALVTLLGGDKALSSINREKMSCFLRMKDTSGGFRHMDMGEMDVR 210
```


